STRUCTURE AND FUNCTION OF HUMAN ERYTHROCYTE PYRUVATE KINASE: MOLECULAR BASIS OF NONSPHEROCYTIC HEMOLYTIC ANEMIA

Giovanna Valentini‡,†*, Laurent R. Chiarelli‖, Riccardo Fortin‡, Manuela Dolzan‡, Alessandro Galizzi‡, Donald J. Abraham§, Changqing Wang§, Paola Bianchi‖, Alberto Zanella‖ and Andrea Mattevi‡*

‡ Dipartimento di Genetica e Microbiologia, Università di Pavia, via Abbiategrasso 207, 27100 Pavia, Italy

‖ Dipartimento di Biochimica, Università di Pavia, via Taramelli 3b, 27100 Pavia, Italy

§ Department of Medicinal Chemistry, Virginia Commonwealth University, 800 East Leigh St. Suite 212 Richmond, VA 23219, USA

‖ Divisione di Ematologia, IRCCS Ospedale Maggiore di Milano, via Francesco Sforza 35, 20122 Milano, Italy

*Corresponding authors: Andrea Mattevi or Giovanna Valentini, Dipartimento di Genetica e Microbiologia, Università di Pavia, via Abbiategrasso 207, 27100 Pavia, Italy; E-mail mattevi@ipvgen.unipv.it, Fax +39-0382-528496

Running title: Human erythrocyte pyruvate kinase.

Copyright 2002 by The American Society for Biochemistry and Molecular Biology, Inc.
Summary

Deficiency of human erythrocyte R isozyme (RPK) is, together with glucose 6-phosphate dehydrogenase deficiency, the most common cause of the nonspherocytic hemolytic anemia. To provide a molecular framework to the disease, we have solved the 2.7 Å resolution crystal structure of human RPK in complex with fructose 1,6-bisphosphate, the allosteric activator, and phosphoglycolate, a substrate analogue, and we have functionally and structurally characterized eight mutants (G332S, G364D, T384M, D390N, R479H, R486W, R504L, R532W) found in RPK-deficient patients. The mutations target distinct regions of RPK structure, including domain interfaces and catalytic and allosteric sites. The mutations affect to different extent thermostability, catalytic efficiency and regulatory properties. These studies are the first to correlate the clinical symptoms with the molecular properties of the mutant enzymes. Mutations greatly impairing thermostability and/or activity are associated to severe anemia. Some mutant proteins exhibit moderate changes in the kinetic parameters, which are sufficient to cause mild-to-severe anemia, underlining the crucial role of RPK for erythrocyte metabolism. Prediction of the effects of mutations is difficult since there is no relation between the nature and location of the replaced amino acid and the type of molecular perturbation.
Characterization of mutant proteins may serve as a valuable tool to assist with diagnosis and genetic counseling.
**Introduction**

Pyruvate kinase (PK)\(^1\) catalyses the conversion of phosphoenolpyruvate (PEP) to pyruvate with the synthesis of ATP. The enzyme requires \(K^+\) and \(Mg^{++}\) (or \(Mn^{++}\)) for activity \((1,2,3)\). The PK catalyzed reaction represents the last step of glycolysis with the reaction product, pyruvate, being involved in a number of energetic and biosynthetic pathways. PK is activated homotropically by PEP and heterotropically by monophosphorylated or bisphosphorylated sugars \((2)\). In addition, \(Mg^{++}\), \(H^+\) and other cations modulate enzymatic activity \((4)\). The regulatory behavior of PK varies depending on the enzyme source. Four PK isozymes have been identified in mammals \((5)\). The M1 (muscle) and M2 (fetal) proteins are products of the alternative splicing of the same mRNA. M2 PK is allosterically activated by fructose 1,6-bisphosphate (FBP) and PEP while the M1 enzyme is exceptional in that it is the only known PK that displays hyperbolic kinetics. The other two mammalian PK isozymes, liver (L) and erythrocyte (R), are coded by the same \(PKLR\) gene through the use of tissue specific alternate promoters. Both R and L isozymes are activated by PEP and FBP \((2)\).

The three-dimensional structures of several PKs from prokaryotic and eukaryotic organisms have been elucidated \((6,7,8,9,10)\). They reveal a conserved architecture. PK is a 200 kDa tetramer with four

---

\(^1\)PK catalyzes the conversion of PEP to pyruvate and ATP.
identical subunits, each consisting of four domains (Fig. 1); the small N-terminal helical domain (absent in bacterial PKs), the A domain with \((\beta/\alpha)_8\) barrel topology, the B domain which is inserted between strand \(\beta 3\) and helix \(\alpha 3\) of the A domain \((\beta/\alpha)_8\) barrel, and the C domain with a \(\alpha+\beta\) topology. This multidomain architecture is instrumental to the regulation of PK activity. The enzyme activation is thought to involve a combination of domain and subunit rotations, coupled to alterations in the active site geometry. In this mechanism, the residues located at the domain and subunit interfaces are crucial in that they function in the communication between activator-binding site and catalytic center \((8,9,10)\).

Deficiency of human erythrocyte R isozyme (RPK) is, together with glucose 6-phosphate dehydrogenase deficiency, the most common cause of nonspherocytic hemolytic anemia \((11)\). RPK deficiency severely affects the erythrocyte metabolism, causing ATP depletion which ultimately leads to hemolysis. Worldwide, more than 150 mutations in the gene coding RPK have been found in RPK-deficient patients \((12)\). The disease is transmitted as recessive trait and the pathological symptoms occur only in homozygotes or compound heterozygotes. The clinical manifestations vary from mild to severe anemia, which can be life threatening and require continuous transfusion therapy. Here, we describe the first crystal structure of
recombinant RPK and the biochemical characterization of eight mutants found in patients subjected to clinical follow-up. These studies allow a correlation between the clinical symptoms and the molecular properties of the mutant enzymes.

**Experimental procedures**

**Expression vectors**

The vectors used to express RPK and its mutant and truncated forms were derivatives of pTrcHisB (Invitrogen). RPK cDNA insert was obtained from pGG1 (13) after introduction of NcoI-NdeI sites around the ATG initiation codon. The mutagenic primer (14) used to introduce the two sites into pGG1 was 5’-CAAGGAGGCTGAAACCATGGCTAGCCAGGAGAACATATCATT. It altered the second and the third codon of the insert. The underlined sequences indicate the mutated bases. The selection primer used to abolish the vector unique restriction site AflIII was 5’-CGCAGGAAAGAACCCTTGGGAGCAAAAAGGCC. The pTrcHisB with RPK cDNA inserted in NcoI/EcoRI and designated pCW3 was mutagenised to restore the codons previously changed. The mutagenic primer was 5’-TAAGGAGGAATAAACCATGTCATCCA GGAGAACATATCAT. The selection was performed by digesting the parental pCW3 with
The new plasmid, containing the correct insert, was named pLC1. To obtain the desired RPK mutants, the pLC1 was subjected to site-directed mutagenesis (14). The same selection primer (5’-CCCCCCTGAATTCGAACTTT GGCTG) to abolish the unique HindIII site was used in all cases. The specific mutagenic primers were:

5’-CTGGAGGTGAGCGACAGCATCATGTTGGCA for G332S;
5’-CTGCAACTTGGCCGGACAAGGCTGTTGTCTG for G364D;
5’-CAAGCCCCGGGCATAAGGGGCGAGACAAG for T384M;
5’-AGGAGAGACAGCAAATGTCGCATAAGGCTG for D390N;
5’-CTGACCACTCTGCGCACTACAGCCCCAGCTTCTG for R479H;
5’-AGCCCGACTCTTCTTGCACTCGGCCTTCA for R486W;
5’-CTGCCACGGCTGGCCTTATCTCCAGGTCACCTTAT for R504L;
5’-ATGATGTAAGATCGCTGGGTGCAATTTGGCA for R532W.

To obtain a truncated form of RPK, lacking the first 49 residues, pCW3 was mutagenised by using the primer 5’-TAAGGAGGAATAAAAGCATTGGAGCTGGGCACTGCC TTCTTCC. This sequence corresponds to that of the plasmid upstream of the initiation triplet ATG and continues with that of the RPK insert starting from the GAG codon of Glu at position 50. The selection was performed by digesting pCW3 with NheI restriction enzyme. The plasmid with the insert encoding the truncated RPK (50-574) was designated pLC3. All inserts were sequenced.
Protein purification and enzymatic analysis

*E. coli* DH5α transformed with the specific expression vectors were grown at 37°C in Luria-Bertani medium containing 100 µg/mL ampicillin. When the culture optical density at 600 nm reached a value of 0.5, the expression was induced by addition of isopropyl-β-D-thiogalactopyranoside at a final concentration of 0.5 mM. The induction time was 12 hours while the induction temperature was 30°C, with the exception of the mutants G332S, G364D, R504L, and R532W, for which the induction temperature was 21°C. Wild-type and mutant enzymes were purified by the procedure of Wang *et al.* (13). Enzyme activities were measured at 37°C by the assay (13) recommended by the International Committee for Standardization in Hematology. Kinetic parameters were determined with the Enzyme Kinetic Module™ 1.1 (SPSS Science Software Gmb). Thermal stability was measured by incubating the enzyme (100-200 µg/mL) at 53°C in a solution consisting of 20 mM potassium phosphate pH 6.5, 100 mM KCl, and 1 mM EDTA. Samples were removed at intervals and immediately assayed.
Crystallography

Recombinant wild-type RPK was crystallized using the vapor diffusion method at 22°C. Well solutions consisted of 50 mM Mes/KOH pH 6.4, 10 mM MnSO₄, and 10-14% w/v PEG8000. Hanging drops were formed by mixing equal volumes of 12 mg/ml protein in 50 mM KCl, 5 mM FBP, 5 mM phosphoglycolate, 20 mM potassium phosphate pH 7.0 and well solutions. The crystals were difficult to reproduce. The recombinant enzyme undergoes partial proteolysis and about 50% of the purified protein chains lack the first 47 amino acids (13). On this basis, we produced a mutant truncated enzyme lacking the first 49 residues (see above). Employment of the truncated protein greatly improved the crystallization, which was carried out using the above-described protocol. Crystals were obtained for the T384M, R479H and R486W truncated mutants by the same protocol used for the truncated wild-type RPK.

RPK crystals belong to space group P2₁ with one tetramer in the asymmetric unit. Diffraction data were measured at 100 K on beamline ID14-EH2 of the European Synchrotron Radiation Facility (Grenoble, France) using a MarCCD detector and beamline B7WB of DESY/EMBL (Hamburg, Germany) using a Mar Imaging Plate. Data processing and reduction were carried out using MOSFLM (15) and
programs of the CCP4 suite (16). Data collection statistics are reported in Table 1. The structure of the wild-type RPK was solved by molecular replacement using the program Molrep (16). The search model was the structure of rabbit muscle M1 PK in complex with pyruvate (7; PDB entry 1PKN). Phases were improved by 4-fold averaging (17) producing an electron density of excellent quality. Model building was carried out with the program O (18). The model was refined using Refmac (19). All measured data (no \( \sigma \) cut-off) were employed and 2.5% of unique reflections were used to monitor the progress of the refinement by \( R_{\text{free}} \) validation. The refined wild-type coordinates provided the starting model for the refinement of the mutants. The set of reflections for calculation of \( R_{\text{free}} \) was identical to that of the wild-type structure refinement. A summary of refinement statistics is presented in Table 1. Analysis and inspection of the structures were carried out with O (18) and programs of the CCP4 package (16). Figures were generated with Molscript (20).
Results

Position of Figure 1

The three-dimensional structure of RPK

The crystallographic studies were performed using a truncated RPK in which the 49 N-terminal amino acids are absent. Use of the truncated protein resulted in considerable improvement in the reproducibility of the crystallization experiments. The truncated protein exhibits kinetic properties virtually identical to those of wild-type RPK. A more detailed analysis of this and other mutants targeting the N-terminal residues will be published elsewhere.

The truncated recombinant RPK was crystallized in complex with phosphoglycolate (a PEP analogue), FBP, Mn++ and K+. The presence of the allosteric activators implies that the crystalline enzyme is in the active R-state. The 2.7 Å resolution structure of RPK reveals the typical four-domain subunit architecture found in all PKs of known three-dimensional structure (Fig. 1a). The A (residues 85-159 and 263-431) and C domains (432-574) together with the small N-terminal domain (57-84) form the main body of the subunit. The B domain (160-262) is loosely packed to the rest of the molecule and adopts slight different orientations (about 4°) in the four
crystallographically independent polypeptide chains. The four subunits of the RPK tetramer are assembled to form a $D_2$ symmetric oligomer. The intersubunit interactions define two large contact areas; the A/A’ interface involves the A domains of subunits related by the vertical twofold axis, as defined in Figure 1b, while the C/C’ interface involves the C-domains of subunits interacting along the horizontal axis.

The structure of RPK subunit closely resembles that of rabbit muscle M1 PK, as expected from the 59% sequence identity between the two proteins. The similarity is highest with M1 PK in complex with pyruvate (7), with a root-mean-square difference of 1.2 Å for 512 Cα atom pairs. This M1 PK complex exhibits the same B domain orientation found in RPK. In other M1 structures, the B domain is either more open, as in the phospholactate complex (21), or more closed, as in the complex with ATP (22,23).

Position of Figure 2

The allosteric site and the catalytic center

RPK was cocrystallized with FBP, phosphoglycolate and the K+ and Mn++ ions. All ligands are clearly visible in the electron density map.
Phosphoglycolate, a potent PK inhibitor (9), is positioned in the PEP-binding site, which is located at the top of the A domain \((\beta/\alpha)_8\) barrel, facing a cleft between the A and B domains (Fig. 1a). It is at the heart of an intricate network of H-bonds, which involve protein residues and the Mn\(^{++}\) and K\(^+\) cations (Fig. 2a). The phosphate group is bound to the K\(^+\) atom and the side chain of Arg116 while the carboxylate moiety is anchored through interactions with the Mn\(^{++}\) ion, the side chain of Thr371 and the main chain nitrogen atoms of Gly338 and Asp339, which are located at the N-terminus of a short helical segment belonging to loop 6 of the A domain \((\beta/\alpha)_8\) barrel. This binding geometry is identical to that observed in the yeast PK-phosphoglycolate complex (9) and closely resembles the binding of pyruvate and phospholactate to rabbit M1 PK (for a discussion of the implications of this binding mode for catalysis see References 7 and 21). These similarities are in keeping with the strict conservation among the PK sequences of all residues surrounding the substrate-binding site.

The FBP activator is hosted in the allosteric site in the C domain (Fig. 1a). The ligand is sandwiched between loops 475-479 and 557-566 (Fig. 2b) and extensively interacts with protein. The 6’-phosphate group is engaged in a salt bridge with Arg532 while the 1’-phosphate is H-bonded to the side chains of Thr475 and Ser480 and
the backbone nitrogen atoms of Thr476 and Thr477. Moreover, the
dipole of \( \alpha \)-helix 480-486 points towards the 1’-phosphate, further
compensating the ligand negative charge. This geometry in FBP-
binding is identical to that found in yeast PK crystallized in the R-
state (9).

Rational for the mutagenesis studies

A survey of the missense mutations associated to the
nonspherocytic hemolytic anemia shows that most of them cluster
in specific regions of the protein three-dimensional structure: the
interface between the A and C domains, the A/A’ intersubunit
interface, the hydrophobic core of the A domain, and the FBP-
binding site (24). We generated eight RPK mutants (Fig. 1a),
targeting residues belonging to each of these regions of the protein.
Almost all selected mutations have been found in homozygote
patients. The kinetic, allosteric and thermostability parameters of
mutant proteins were evaluated (Table 2) and the crystal structures
of three mutants (T384M, R479H and R486W) were solved. The
mutations did not induce significant conformational changes in the
overall protein conformation and, therefore, we shall restrict the
description of the mutant structures mainly to the sites affected by
the mutations.
**Position of Figure 3**

**G332S mutation in the A domain hydrophobic core**

Many RPK mutations are localized in the hydrophobic core of the A domain. An example is the G332S mutation (nucleotide 994G->A), which affects a residue that is strictly conserved among PK sequences. Gly332 is located on strand $\beta 6$, being buried inside the domain core (Fig. 1a). The G332S protein exhibits a 9-fold decrease in the catalytic efficiency (5-fold for the FBP activated protein) and is considerably less thermostable than the wild-type enzyme (Table 2 and Fig. 3). These substantially altered molecular properties account for the clinical data. In homozygous form, the G332S mutation leads to a severe anemia with the need of regular transfusions (25,26).

**Mutations at the A/C interface: G364D, R486W and R504L**

The interface between the A and C domains is characterized by many polar interactions that often involve charged side chains. Many of these residues represent sites of pathological mutations, which cause RPK-deficiency with variable levels of severity.
The mutation R504L (nucleotide 1511G->T) affects Arg504, a C-domain residue which is partly solvent accessible and engaged in an interdomain salt bridge with Asp281 (Fig. 1a). The R504L mutation removes this interdomain interaction and introduces a hydrophobic Leu side chain in a solvent exposed site close to a negatively charged Asp. Such amino acid replacement is clearly unfavorable, providing a rational for the extreme instability of the protein, which prevented functional analysis (Table 2). This feature explains the severe anemia found in RPK-deficient patients homozygous for this mutation (27).

position of figure 4

The other two investigated mutants targeting the A/C interface affect Gly364 and Arg486, which are part of a region of close association between the A and C domains (Figs. 1a and 4a). Arg486 is H-bonded to the carbonyl oxygen of Leu362 at the C-terminus of the A domain helix α6 while the neighboring Gly364 allows a sharp turn of the polypeptide chain with a backbone conformation (φ=85°,ψ=98°) that is unfavorable for a non-glycine residue. The G364D (nucleotide 1091G->A) mutation has a drastic effect on the enzyme stability, which is coupled to a 3-fold reduction of the
catalytic efficiency (Fig. 3b and Table 2). Given the tightly packed environment and the backbone conformation of Gly364, it is conceivable that introduction of a charged Asp side chain at this site of the A/C interface can greatly perturb the domain assembly, thus being deleterious for stability. Fully consistent with these observations is the severe anemia found in patients homozygous for G364D (28). Together with R504L, the G364D mutant highlights the notion that the inter-domain interactions at the A/C interface are critical for the stability of the protein.

R486W (nucleotide 1456C->T) is among the most frequent mutations found in RPK-deficient patients (11). Characterization of this mutant reveals that such a drastic amino acid replacement results in small effects on the molecular properties. The mutant three-dimensional structure shows that the Trp side chain is accommodated without any structural perturbation. With respect to the wild-type structure, no atomic movements larger than 0.25 Å can be detected whereas the indole nitrogen atom is able to establish a H-bond with the carbonyl oxygen atoms of Leu362 and Asp361. Such structural conservation matches the limited changes in biochemical parameters. The thermostability is even higher than that of the wild-type protein (Fig. 3b) and the allosteric properties are essentially unmodified (Table 2 and Fig. 3a). The only significant
perturbation is in the catalytic efficiency, which drops to 30% of the value measured for the wild-type RPK (Table 2). These moderate variations in the molecular parameters correlate with the clinical symptoms since patients homozygous for the R486W mutation generally exhibit a mild anemia (25).

The perturbed kinetics of the R486W protein is puzzling since Arg486 is >20 Å away from the catalytic center (Fig. 1a), which is left unperturbed by the mutation as shown by the mutant crystal structure. The “long range” effect exercised by the R486W mutation might reflect altered dynamic properties. It is known that the B domain adopts different conformations depending on ligand binding (21,22,23). The introduction of the Trp aromatic ring may restrict the overall ability of the enzyme to undergo the conformational changes occurring during catalysis, thereby perturbing the reaction kinetics.

The A/A’ interface: T384M and D390N

Asp390 is a solvent inaccessible residue located in the A/A’ interface, at the heart of a H-bond network that involves Arg337 and Ser389’ (the prime symbol denotes a residue of a different subunit). Based on the comparison between the structures of the T-
state *E. coli* PK and of the M1 isozyme, it was found that Asp390 is crucial for the allosteric transition by coupling changes in the quaternary structure with alterations in the active site geometry (8). A pathological mutation affecting this residue (D390N corresponding to nucleotide 1168G->A) has been detected in a heterozygote patient (25). The molecular analysis shows that the D390N amino replacement causes the almost complete inactivation of the protein, which, however, is not less thermostable than the wild-type RPK (Table 1 and Fig. 3b). These results are very similar to those obtained with the *E. coli* enzyme for which the same mutation was investigated (29). These observations support the idea that Asp390 has a key role in the allosteric regulation, suggesting that the D390N mutation may lock the protein in an inactive conformation, impairing the transition to the R-state.

**Position of Figure 5**

The mutation T384M (nucleotide 1151C->T) affects a residue, which, though not directly involved in intersubunit interactions, lays very close to the A/A’ molecular twofold axis (Fig. 5). Thr384 is located at the N-terminus of helix α7 of the A domain (β/α)$_8$ barrel. Its OG atom is H-bonded to the backbone nitrogen atoms of Ala386 and Glu387, thus acting as helix-capping element. The three-
dimensional structure of the T384M mutant reveals that the mutation does not cause atomic shifts larger than 0.3 Å (Fig. 5). The bulkier Met side chain is easily accommodated, the only change being the removal of the helix-capping H-bonds. Also the kinetic characterization shows limited variations, the main difference between T384M and the wild-type protein being a 3-fold reduction of the catalytic efficiency (5-fold for the FBP activated form; Table 2) mainly accounted by a reduction in $K_{cat}$. Likewise, the mutation does not alter the thermostability parameters (Fig. 3).

Thr384 is not part of the binding sites for PEP and ADP and the crystal structure of the T384M protein shows that the active site geometry is not affected by the mutation. Thus, the altered kinetics displayed by the T384M mutant is difficult to rationalize. Modified enzymatic parameters were observed also for the equivalent mutation in the rabbit kidney isozyme (30). Thr384 is close to the contact region between A and B domains (Fig. 1a) and, therefore, may disturb the “closure” of the B domain occurring on ATP binding (22). An alteration of the equilibrium between “open” and “closed” B domain conformations may affect the enzymatic activity. It is remarkable that homozygosity for the T384M mutation is associated to anemia with mild-to-severe symptoms (31,32),
implying that even moderate changes in the enzyme catalytic power can have pathological effects.

*The allosteric site: R532W and R479H*

The negative charges of FBP are compensated by the N-terminus of helix 479-486 for the 1’-phosphate and Arg532 side chain for the 6’-phosphate (Fig. 2b). We have investigated two mutations that target both these elements involved in FBP binding. The first of these mutations is R532W (nucleotide 1594 C->T) that has been found in compound heterozygotes, in which the other allele had a mutation causing the truncation of the protein. The clinical symptoms in the patients carrying the mutation were severe (33). Molecular analysis of R532W protein indicates a complete loss in the responsiveness to FBP, highlighting the essential of Arg532 role for activator binding (Table 2). These perturbed allosteric properties are associated to a decreased thermostability (Fig. 3b), possibly reflecting the energetically unfavorable exposure on the protein surface of the hydrophobic Trp residue.

The R479H mutation has been found in RPK-deficient patients affected by severe anemia (34,35). The side chain of Arg479 is located in the neighborhood of FBP although it does not directly
interact with the activator (Fig. 2b). The crystal structure of R479H is identical to that of the wild-type protein, with the His side chain being fully solvent-exposed. Similarly, the kinetic parameters (Table 2) appear to be essentially unaffected by the mutation. These features are in contrast with the severe clinical symptoms (34,35). An explanation for this riddle is given by the observation that the mutation affects nucleotide 1436, which is located on a splicing site at the 3'-end of exon 10. This fact together with our biochemical analysis suggests that, rather than the amino acid replacement, defects in mRNA splicing process are the actual cause of the RPK-deficiency.

Discussion

Implications for the allosteric regulation

PK is a typical allosteric enzyme of the K-type. The allosteric signal is transmitted across the long distance (>20 Å) separating the FBP-binding site from the catalytic center. The exact mechanism of the allosteric transition is not known in detail since no PK has been so far crystallized in both T and R states. The comparison between the structures of the T-state E. coli PK and of the rabbit M1 enzyme (8)
suggested that the allosteric transition involves modifications in the relative orientations of the domains and subunits coupled to conformational changes in the PEP-binding site. The X-ray analysis of the T-state *Leishmania mexicana* PK (10) and the R-state yeast PK (9) led to a further refinement of this model, allowing to discriminate between the structural differences that are consequence of the inherent divergence between eukaryotic and prokaryotic proteins and the conformational changes that are genuinely due to the allosteric transition. The mechanism of PK regulation has also been the subject of many mutagenesis experiments (23 and 36 and references therein). The general picture emerging from the mutant analysis is that the intersubunit interactions at the A/A’ and C/C’ interfaces and the interdomain interactions at the A/B interface are key to the allosteric responsiveness and to define the distribution of the conformations between active and inactive states. Moreover, the mutagenesis analyses combined with the crystallographic data provide clear evidence for the idea that the T and R forms correspond to ensembles of conformations characterized by rotational flexibility of the B-domain (23).

Our study on human RPK is fully consistent with these features. The crystal structure shows that also in RPK the B-domain is flexible, adopting different orientations in the crystallographically
independent subunits. The key functional role of the A/A’ interface is highlighted by D390N mutation, which targets a residue located in the core of the A/A’ interface, producing an enzyme that retains a stable tetrameric state but almost entirely lacks enzymatic activity. Conversely, none of the mutations targeting residues at the A/C interface alters the enzyme allostERIC properties. Thus, in agreement with the recent mutagenesis data on the yeast PK (36), this domain interface appears to have little role in the transduction of the allostERIC signal, rather being important for the stability of the domain assembly within the enzyme subunit.

Molecular basis of nonspherocytic hemolytic anemia

Characterization of mutant proteins shows that amino acid substitutions can affect thermostability, catalytic efficiency and response to the allostERIC effector. Various regions of the RPK structure, including domain interfaces and functional sites, are affected by the pathological mutations (Fig. 1a). However, there appears to be no relation between the nature and location of the replaced amino acid and the type of molecular perturbation. For instance, both R504L and R486W mutations affect Arg residues involved in interdomain polar interactions at the A/C interface but their effects are substantially different. The R504L protein is
extremely unstable while the R486W mutant is even slightly more thermoresistant than the wild-type (Fig. 3b). These observations emphasize the difficulty of predicting the consequences of mutations simply from the location and the nature of the target residues. They also warn against predictions of the effects of mutations in human RPK based on the molecular analysis of other PK isoenzymes.

The clinical manifestations of a genetic disease reflect the interactions of a variety of physiological and environmental factors and do not solely depend on molecular properties of the altered molecule. Given this caution, it is evident that there is a general correlation between the clinical manifestations and the biochemical parameters of the mutant proteins. Mutants exhibiting strongly perturbed kinetic and thermostability parameters (G332S, G364D, R504L and R532W) are associated to severe RPK-deficiency. Conversely, in the case of less abnormal molecular properties, the disease has milder manifestations. It is remarkable that pathological conditions are present in association to mutations such as T384M or R486W, which are simply characterized by a moderate reduction of the catalytic efficiency. The physiological concentrations of RPK substrates and effectors are in the \( \mu \text{M} \) range (37). Therefore, in vivo RPK operates in sub-saturating conditions, which may amplify the
effects of the different catalytic efficiencies between the wild-type and mutant proteins (Fig. 3a).

In conclusion, our studies indicate that the functional parameters of RPK are so finely tuned that even moderate molecular alterations may significantly perturb cell metabolism. The correlation between molecular and clinical parameters in PK-deficiency suggests that biochemical characterization of mutant proteins may serve as a valuable tool to understand and assist with diagnosis and genetic counseling.

Acknowledgements

We thank the staff of DESY/EMBL and ESRF for help during data collection. This research was supported by grants from University of Pavia (Progetto d’Ateneo “Nuove tecnologie molecolari e cellulari”), Ministero della Ricerca Scientifica e Tecnologica (Progetto Genomica Funzionale), and Allos Therapeutics Inc., Westminster, CO (USA).
References


**Footnotes**

1 Abbreviations: PK, pyruvate kinase; RPK, human erythrocyte pyruvate kinase; PEP, phosphoenolpyruvate; FBP, fructose 1,6-bisphosphate.

2 PDB deposition: Atomic coordinates and structure factors have been deposited in the Protein Data Bank (accession codes xxx and xxx for native RPK, xxx and xxx for T384M, xxx and xxx for R479H, and xxx and xxx for R486W).
Table 1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>Native</th>
<th>R479H</th>
<th>T384M</th>
<th>R486W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2₁</td>
<td>P2₁</td>
<td>P2₁</td>
<td>P2₁</td>
</tr>
<tr>
<td>Unit cell axes a,b,c, Å</td>
<td>74.4</td>
<td>74.0</td>
<td>76.3</td>
<td>73.7</td>
</tr>
<tr>
<td></td>
<td>172.1</td>
<td>171.8</td>
<td>173.0</td>
<td>171.2</td>
</tr>
<tr>
<td></td>
<td>85.5</td>
<td>85.1</td>
<td>85.8</td>
<td>85.0</td>
</tr>
<tr>
<td>β, °</td>
<td>92.5</td>
<td>91.2</td>
<td>93.1</td>
<td>91.6</td>
</tr>
<tr>
<td>Completenessa, %</td>
<td>95.9</td>
<td>92.6</td>
<td>93.7</td>
<td>89.0</td>
</tr>
<tr>
<td></td>
<td>(86.5)</td>
<td>(90.1)</td>
<td>(91.3)</td>
<td>(68.6)</td>
</tr>
<tr>
<td>Measured reflections</td>
<td>118752</td>
<td>98901</td>
<td>116305</td>
<td>76390</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>55091</td>
<td>50701</td>
<td>52721</td>
<td>42512</td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>2.7</td>
<td>2.7</td>
<td>2.7</td>
<td>2.9</td>
</tr>
<tr>
<td>Rsyma,b , %</td>
<td>7.6</td>
<td>9.5</td>
<td>7.9</td>
<td>10.3</td>
</tr>
<tr>
<td></td>
<td>(24.1)</td>
<td>(46.7)</td>
<td>(54.6)</td>
<td>(28.3)</td>
</tr>
<tr>
<td>Protein atoms</td>
<td>15431</td>
<td>15181</td>
<td>15435</td>
<td>15218</td>
</tr>
<tr>
<td>Solvent molecules</td>
<td>65</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ligand atoms</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>124</td>
</tr>
<tr>
<td>R-factorc, %</td>
<td>23.0</td>
<td>24.2</td>
<td>24.6</td>
<td>24.3</td>
</tr>
<tr>
<td>Rfree c, %</td>
<td>27.9</td>
<td>29.0</td>
<td>30.0</td>
<td>30.9</td>
</tr>
<tr>
<td>rms bond lengths, Å</td>
<td>0.020</td>
<td>0.025</td>
<td>0.024</td>
<td>0.024</td>
</tr>
<tr>
<td>rms bond angles, °</td>
<td>2.0</td>
<td>2.4</td>
<td>2.8</td>
<td>2.9</td>
</tr>
<tr>
<td>NCSd, domains A and C, Å</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>subunit 1 – subunit 2</td>
<td>0.26</td>
<td>0.15</td>
<td>0.22</td>
<td>0.17</td>
</tr>
<tr>
<td>subunit 1 – subunit 3</td>
<td>0.16</td>
<td>0.27</td>
<td>0.19</td>
<td>0.45</td>
</tr>
<tr>
<td>subunit 1 – subunit 4</td>
<td>0.20</td>
<td>0.24</td>
<td>0.22</td>
<td>0.28</td>
</tr>
<tr>
<td>NCSd, domain B, Å</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>subunit 1 – subunit 2</td>
<td>0.37</td>
<td>0.32</td>
<td>0.31</td>
<td>0.26</td>
</tr>
<tr>
<td>subunit 1 – subunit 3</td>
<td>0.33</td>
<td>0.40</td>
<td>0.32</td>
<td>0.33</td>
</tr>
<tr>
<td>subunit 1 – subunit 4</td>
<td>0.31</td>
<td>0.37</td>
<td>0.43</td>
<td>0.38</td>
</tr>
<tr>
<td>Ramachandran plotc, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most favoured regions</td>
<td>89.9</td>
<td>86.6</td>
<td>86.7</td>
<td>83.2</td>
</tr>
<tr>
<td>Additionally allowed regions</td>
<td>9.5</td>
<td>12.4</td>
<td>12.6</td>
<td>15.6</td>
</tr>
<tr>
<td>Disallowed regions</td>
<td>0.5</td>
<td>0.9</td>
<td>0.7</td>
<td>1.2</td>
</tr>
</tbody>
</table>

aValues in parentheses are for reflections in the highest resolution shell.

bRsym = Σ|Ii - <I>|/ΣIi, where Ii is the intensity of ith observation and <I> is the mean intensity of the reflection.
\[ R\text{-factor} = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|} \] where \( F_{\text{obs}} \) and \( F_{\text{calc}} \) are the observed and calculated structure factor amplitudes, respectively.

Root-mean-square deviation between \( C_\alpha \) atoms of the non-crystallographically symmetry related monomers present in the asymmetric unit. Tight NCS restraints were applied throughout the refinement.

\( \text{Analyzed with Procheck (38).} \)
Table 2. Kinetic parameters of the wild-type and mutant RPKs

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$K_{cat}$ (s$^{-1}$)</th>
<th>$S_{0.5}$ (mM)</th>
<th>$k_{cat}/S_{0.5}$ (s$^{-1}$/mM)</th>
<th>$n_H$</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$S_{0.5}$ (mM)</th>
<th>$k_{cat}/S_{0.5}$ (s$^{-1}$/mM)</th>
<th>$n_H$</th>
<th>$K_{cat}$ (s$^{-1}$)</th>
<th>$S_{0.5}$ (mM)</th>
<th>$k_{cat}/S_{0.5}$ (s$^{-1}$/mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type$^c$</td>
<td>355±12</td>
<td>1.10±0.04</td>
<td>323</td>
<td>1.60±0.16</td>
<td>355±11</td>
<td>0.18±0.020</td>
<td>1972</td>
<td>1.05±0.07</td>
<td>355±13</td>
<td>0.17±0.01</td>
<td>2080</td>
</tr>
<tr>
<td>G332S</td>
<td>137±6</td>
<td>3.79±0.2</td>
<td>36</td>
<td>2.31±0.12</td>
<td>152±5</td>
<td>0.38±0.04</td>
<td>389</td>
<td>1.08±0.08</td>
<td>111±7</td>
<td>0.49±0.03</td>
<td>226</td>
</tr>
<tr>
<td>G364D</td>
<td>104±7</td>
<td>0.93±0.03</td>
<td>112</td>
<td>1.54±0.03</td>
<td>118±6</td>
<td>0.75±0.02</td>
<td>153</td>
<td>1.39±0.03</td>
<td>115±6</td>
<td>0.16±0.04</td>
<td>718</td>
</tr>
<tr>
<td>T384M</td>
<td>149±10</td>
<td>1.24±0.09</td>
<td>120</td>
<td>1.50±0.03</td>
<td>172±7</td>
<td>0.36±0.07</td>
<td>383</td>
<td>1.00±0.16</td>
<td>135±9</td>
<td>0.15±0.02</td>
<td>900</td>
</tr>
<tr>
<td>D390N</td>
<td>0.48±0.04</td>
<td>1.40±0.05</td>
<td>0.34</td>
<td>1.65±0.01</td>
<td>0.55±0.05</td>
<td>0.34±0.009</td>
<td>1.6</td>
<td>1.15±0.02</td>
<td>0.45±12</td>
<td>0.25±0.01</td>
<td>1.8</td>
</tr>
<tr>
<td>R479H</td>
<td>390±8</td>
<td>1.10±0.03</td>
<td>454</td>
<td>2.09±0.02</td>
<td>386±8</td>
<td>0.08±0.003</td>
<td>4452</td>
<td>1.17±0.06</td>
<td>381±12</td>
<td>0.20±0.02</td>
<td>1905</td>
</tr>
<tr>
<td>R486W</td>
<td>195±4</td>
<td>1.69±0.06</td>
<td>116</td>
<td>2.07±0.11</td>
<td>203±10</td>
<td>0.40±0.07</td>
<td>492</td>
<td>1.32±0.03</td>
<td>218±13</td>
<td>0.24±0.02</td>
<td>908</td>
</tr>
<tr>
<td>R504L</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>R532W</td>
<td>183±5</td>
<td>0.63±0.03</td>
<td>290</td>
<td>1.41±0.11</td>
<td>187±8</td>
<td>0.66±0.06</td>
<td>275</td>
<td>1.48±0.14</td>
<td>189±12</td>
<td>0.20±0.04</td>
<td>945</td>
</tr>
</tbody>
</table>
Results are means (SE) for 3 determinations from 4 different protein preparations

a Kinetic parameters for PEP were obtained at fixed 1.5 mM ADP by fitting data to the Hill plot.

b Kinetic parameters for ADP were obtained at fixed 5 mM PEP by fitting data to the Lineweaver-Burk plot.

c Data from (13).
Figure legends

**Figure 1.** Three-dimensional crystal structure of RPK. The N-terminal domain is yellow, the A domain is red, the B domain is cyan and the C domain is green. (a) The RPK subunit. The gray spheres indicate the Cα atoms of the residues subjected to mutagenesis. (b) The RPK tetramer. In this orientation, a molecular twofold axis is perpendicular to the plane of the paper whereas the other two molecular twofold axes are vertical and horizontal in the pane of the paper (outlined by vertical and horizontal lines, respectively).

**Figure 2.** The allosteric and catalytic sites of RPK. (a) Stereo view of the active site with bound phosphoglycolate (outlined by gray bonds), Mn²⁺ and K⁺. With respect to Figure 1a, the model has been rotated by approximately 30° around an axis horizontal in the plane of the paper. (b) Stereo view of the allosteric site with bound FBP (gray bonds). The orientation is as in Figure 1a.

**Figure 3.** Characterization of RPK mutants. ● wild-type, ○ G332S, ▼ G364D, ▽ T384M, ■ D390N, □ R479H, ◇ R486W, ◆ R532W. (a) Steady-state kinetics of wild-type and mutant RPKs as a function of PEP. (b) Thermal stability of wild-type and mutant RPKs. The residual
activity after incubation at 53°C is expressed as percentage of initial activity.

**Figure 4.** The A/C interface in the region surrounding Arg486. The orientation is as in Figure 1a. (a) Stereo diagram of the wild-type structure. (b) Stereo diagram of the R486W mutant structure.

**Figure 5.** The A/A’ interface close to Thr384. The helices $\alpha7$ of twofold related subunits are shown. Residues of the opposite subunits are denoted by the prime symbol. Superposed to Thr384 is the Met side chain (gray bonds) of the crystal structure of the T384M mutant. The orientation is as in Figure 1a.
Figure 1a
Figure 2a
Figure 2b
Figure 3a
Figure 3b
Figure 4a
Figure 5