

PYK-SubstitutionOME: an integrated database containing allosteric coupling, ligand affinity and mutational, structural, pathological, bioinformatic and computational information about pyruvate kinase isozymes

Liskin Swint-Kruse¹, Larissa L. Dougherty^{1,‡}, Braelyn Page¹, Tiffany Wu¹, Pierce T. O’Neil¹, Charulata B. Prasannan¹, Cody Timmons², Qingling Tang¹, Daniel J. Parente^{1,3}, Shwetha Sreenivasan¹, Todd Holyoak^{1,4} and Aron W. Fenton^{1,*}

¹Department of Biochemistry and Molecular Biology, The University of Kansas Medical Center, 3901 Rainbow Blvd., Kansas City, KS 66160, USA

²Chemistry Department, Southwestern Oklahoma State University, 100 Campus Dr., Weatherford, OK 73096, USA

³Department of Family Medicine and Community Health, The University of Kansas Medical Center, 3901 Rainbow Blvd., Kansas City, KS 66160, USA

⁴Department of Biology, University of Waterloo, 200 University Ave. W, Waterloo, ON N2L 3G1, Canada

*Corresponding author: Tel: (913) 588-7033; Fax: (913) 588-9896; Email: afenton@kumc.edu

‡Present address: Larissa L. Dougherty (larissadougherty@gmail.com; Larissa.L.dougherty.gr@dartmouth.edu), Department of Biochemistry and Cell Biology, Geisel School of Medicine at Dartmouth, Hanover, NH 03755, USA.

Citation details: Swint-Kruse, L., Dougherty, L.L., Page, B. *et al.* PYK-SubstitutionOME: an integrated database containing allosteric coupling, ligand affinity and mutational, structural, pathological, bioinformatic and computational information about pyruvate kinase isozymes. *Database* (2023) Vol. 2023: article ID baad030; DOI: <https://doi.org/10.1093/database/baad030>

Abstract

Interpreting changes in patient genomes, understanding how viruses evolve and engineering novel protein function all depend on accurately predicting the functional outcomes that arise from amino acid substitutions. To that end, the development of first-generation prediction algorithms was guided by historic experimental datasets. However, these datasets were heavily biased toward substitutions at positions that have not changed much throughout evolution (i.e. conserved). Although newer datasets include substitutions at positions that span a range of evolutionary conservation scores, these data are largely derived from assays that agglomerate multiple aspects of function. To facilitate predictions from the foundational chemical properties of proteins, large substitution databases with biochemical characterizations of function are needed. We report here a database derived from mutational, biochemical, bioinformatic, structural, pathological and computational studies of a highly studied protein family—pyruvate kinase (PYK). A centerpiece of this database is the biochemical characterization—including quantitative evaluation of allosteric regulation—of the changes that accompany substitutions at positions that sample the full conservation range observed in the PYK family. We have used these data to facilitate critical advances in the foundational studies of allosteric regulation and protein evolution and as rigorous benchmarks for testing protein predictions. We trust that the collected dataset will be useful for the broader scientific community in the further development of prediction algorithms.

Database URL: <https://github.com/djparente/PYK-DB>

Introduction

To identify important differences among human genomes, to understand the impact from changes that occur as viruses evolve and to engineer novel protein function, scientists need improved predictions about the functional outcomes that result from amino acid substitutions. To that end, numerous algorithms have been developed. The first-generation prediction algorithms incorporated assumptions derived from and tested upon mutational datasets that were generated over several decades of research. However, most of these historical substitution studies were inadvertently biased to positions

that have been conserved throughout evolution. This bias precludes accurate predictions for substitutions at other types of positions: in particular, predictions are poor for a special class of positions—which we have defined as ‘rheostat’ positions (1)—that do ‘not’ follow the rules that are commonly assumed for substitutions at conserved positions (Table 1) (1–10).

More recently, positions with a wider range of conservation (and non-conservation) have been assessed in ‘deep mutational scanning’ experiments (11–14). However, the functional readout for these studies is often indirect (e.g. biological

Received 23 November 2022; Revised 29 March 2023; Accepted 11 April 2023

© The Author(s) 2023. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1. Common assumptions about substitution rules fail for some amino acid positions

Assumptions from historic studies	Refutation
At ‘important’ protein positions	At ‘rheostat positions’
1. Most amino acid substitutions are highly detrimental to structure or function.	The 20 possible substitutions lead to a wide range of functional outcomes. Thus, at a single rheostat position, only a few substitutions are highly detrimental, whereas most exhibit intermediate outcomes and some exhibit enhancing outcomes.
2. Chemically similar side chains can substitute for each other without much detriment.	Chemistries of the substituted amino acids often have little correlation with their functional outcomes. In other words, chemically dissimilar side chains can have more similar functional outcomes than do chemically similar side chains.
3. Substitution outcomes can be extrapolated among homologs.	Functional outcomes from substitutions do ‘not’ extrapolate among homologs.

survival) or of low resolution (e.g. cell sorting). Moreover, these readouts often agglomerate multiple functional parameters; as such, the outcomes associated with different substitutions could arise from different biochemical parameters. For example, mutation outcomes in an enzyme might arise from altered substrate binding, catalytic rate, product binding or allosteric regulation. Deep mutational scanning assays can also be sensitive to protein expression levels, although various strategies can help parse these changes from functional changes (1, 7, 8, 15). As such, although data from deep mutational scanning are useful for training prediction algorithms that are based on machine learning, they do not contain the information needed to predict functional change from fundamental chemical principles (16).

Therefore, our (6, 9, 10, 17–20) and others’ (21–23) approach to addressing the need for improved predictive algorithms has been to create large, biochemical datasets. The database described herein collates the wide range of biochemical, mutational, bioinformatic, structural, pathological and computational studies that have been performed for the homologs of the pyruvate kinase (PYK) isozyme family. The centerpiece of this database is the quantitative, biochemical characterization for more than 1000 variants. Another key feature is that many of the characterized positions were subjected to site-specific, semi-saturating mutagenesis (semiSM; 10–12 substitutions per position).

We reached this semiSM design after considering (and sometimes implementing) several other designs. For example, alanine and glycine scans have been very popular for ‘removing’ the side chain to determine the role of each position (19). However, an alanine variant may function similarly to the native protein even if the substituted position contributes to function (1, 24). Therefore, the functional contributions of such positions may be overlooked if other amino acid substitutions are excluded (24). Furthermore, if only one

amino acid substitution is chosen, results cannot discriminate whether the original amino acid was required for function (a loss-of-function mutation) or the substituted amino acid was simply not tolerated (a negative gain-of-function mutation) (5, 10, 24–26).

These results demonstrate that multiple amino acid substitutions must be assessed for each position. Ideally, 19 substitutions would be created at each position *via* site-saturating mutagenesis (SSM), but this approach can be prohibitively expensive. One approach for simplifying experimental design is to use representative amino acids from chemically similar groups. Unfortunately, this has led to additional bias: as mentioned in Table 1, chemically similar amino acids do not always have similar substitution outcomes at non-conserved positions (1, 3–5, 7, 9, 10, 25, 26).

Fortunately, we have found that the overall functional role of a ‘position’—rather than of a ‘residue’—can be determined from just 10–12 amino acid substitutions (‘semiSM’) (5). To describe each position’s functional role, we previously devised the positional classification system of ‘toggle’, ‘neutral’ and ‘rheostat’ positions (5) that is reviewed in Supplemental Information 1. In brief, (i) at toggle positions, most substitutions abolish function (‘dead’) (2); (ii) at neutral positions, most substitutions function like the wild type (6) and (iii) at rheostat positions, each substitution leads to a different effect on the measured functional parameter, and the set of substitutions samples at least half of the parameter’s accessible range (as defined by wild type and dead) (1, 5). This change of perspective—from the effects of individual substitutions to the overall roles of positions—allows biochemical semiSM results to be readily correlated with bioinformatic and structural data and compared among homologs.

Indeed, the widespread existence of rheostat positions—which do not follow the expected substitution rules (Table 1)—mandates biochemical and biophysical SSM/semiSM studies in order to understand how each amino acid position contributes to a protein’s biochemical properties. Although the first rheostat positions were identified in studies of non-conserved positions (1), later studies showed that properties derived from sequence alignments only partially discriminate rheostat positions from neutral or toggle positions (9, 17). Furthermore, no striking structural changes have been associated with the functional outcomes from substitutions at rheostat positions (8, 18). Thus, more examples must be collected. The current database includes multiple examples of human liver pyruvate kinase (hLPYK) rheostat positions, along with their structural and bioinformatic characteristics.

Overview of PYK function

PYK catalyzes the last step in glycolysis and is present in all domains of life. Humans have four PYK isozymes. hLPYK (expressed in the liver) and hRPYK (expressed in mature erythrocytes) are two products of the *PKLR* (*l/r-pyk*) gene that are generated from alternative start sites. hM₁PYK (PKM1; expressed in the skeletal muscle, heart and brain) and hM₂PYK (PKM2; expressed in all fetal tissues, adult smooth muscle and many cancers) are two gene products from the *PKM* gene that are generated *via* alternative splicing. Among these, hLPYK controls liver homeostasis between glycolysis and gluconeogenesis to regulate blood glucose levels.

In contrast, the PYK isozyme expressed in the skeletal muscle (M₁PYK) appears to have less control over the glycolytic pathway. These and other features of the human isozymes have been extensively reviewed (27–30).

Consistent with their tissue-specific regulatory roles, PYK isozymes have evolved a range of allosteric and post-translational modifications that regulate their enzymatic activities. The predominant regulation in most known PYK isozymes is *via* binding to phosphorylated sugars that allosterically improve affinity (K_{app}) for the substrate phosphoenolpyruvate (‘PEP’). For the flagship isozyme used in many of our studies, hLPYK, affinity for PEP is improved when the allosteric activator fructose-1,6-bisphosphate (abbreviated as ‘Fru-1,6-BP’ in text, ‘FBP’ in subscripts and figures and ‘F16BP’ in the database) is bound; in addition, PEP affinity is reduced when the allosteric inhibitor alanine is bound at yet another allosteric site. Several other PYK isozymes exhibit both activating and inhibiting allosteric regulation. In contrast, a few isozymes, such as *Zymomonas mobilis* PYK (‘ZmPYK’), are ‘not’ allosterically regulated (18, 31–33).

Among the regulated PYK, both allosteric inhibition and activation can be defined by an allosteric energy cycle (Figure 1) (34–38). When characterizing two allosteric functions, as for hLPYK, five functional parameters are quantified (Figure 1): (i) K_{a-PEP} (the kinetic equivalent to K_{ia-PEP}), the apparent affinity for PEP; (ii) K_{ix-Ala} , the binding of Ala; (iii) K_{ix-FBP} , the binding of Fru-1,6-BP; (iv) Q_{ax-Ala} , the allosteric coupling constant that relates PEP and Ala binding to the protein, and (v) Q_{ax-FBP} , the allosteric coupling constant that relates PEP and Fru-1,6-BP binding to the protein. Mathematically, Q_{ax} is defined as follows:

$$Q_{ax} = K_a / -K_{a/x} = K_{ix} / -K_{ix/a}, \quad (1)$$

where K_a is the kinetically derived apparent affinity of the enzyme for PEP in the absence of the effector and $K_{a/x}$ is the affinity of the enzyme for PEP when the effector is present at saturating concentrations. K_{ix} is the binding affinity for the effector in the absence of PEP, and $K_{ix/a}$ is the binding affinity of the effector in the saturating presence of PEP. Note that the effector binding and allosteric coupling constants need not be correlated; indeed, many substitutions at rheostat positions do not show correlated functional changes in those parameters (10).

A key feature of these experiments is that ligand affinity is evaluated in an enzyme concentration-independent measurement. That is, K_a can be determined from $1/2 V_{max}$ without knowing the enzyme concentration. Since allosteric coupling is a ratio of ligand binding values, Q can also be evaluated without knowing the enzyme concentration. This allows for the quantitative evaluation of substrate affinity, allosteric effector binding and allosteric coupling to be completed in partially purified protein samples, thereby facilitating the throughput for a semiSM study. Importantly, the independence from enzyme concentration also ensures that functional data are not ‘contaminated’ with effects on protein stability and provide ‘pure’ information about functional changes. (A limitation is that effects on k_{cat} are not determined in this assay.) Examination of the variants collected in this database illustrate that all five of these functional parameters can be affected by substitutions in hLPYK, with outcomes that range over orders of magnitude (3, 5, 9, 10).

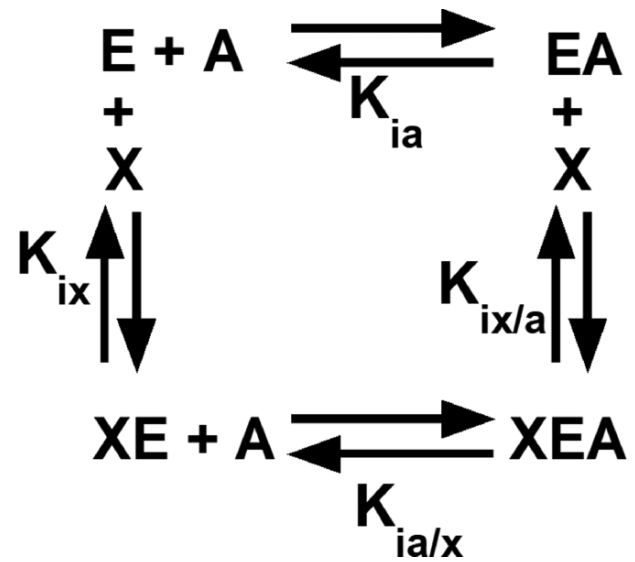


Figure 1. An allosteric energy cycle in which an enzyme (E) binds one substrate (A) and one allosteric effector (X). K_{ia} is the binding of the substrate to the enzyme in the absence of the effector; when initial data are derived from kinetic measurements, this parameter is designated as K_a . $K_{ia/x}$ is the binding of the substrate to the enzyme in the presence of saturating concentrations of the effector. K_{ix} is the binding of the effector to the enzyme when the substrate is absent. $K_{ix/a}$ is the binding of the effector to the enzyme in the presence of saturating concentrations of the substrate. Allosteric coupling is defined as $Q_{ax} = K_a/K_{a/x} = K_{ix}/K_{ix/a}$ (34, 36, 37, 223). A description of hLPYK’s allosteric regulation requires two such functional cycles—one for activation by Fru-1,6-BP and one for inhibition by alanine.

Overview of PYK structure

Many PYK isozymes are homotetramers (27–30) (Figure 2, left); a few PYK isozymes are homodimers (17, 31–33, 39). In most isozymes, each monomer comprises three domains (Figure 2, center). The PYK ‘A’ domain is a TIM barrel (40–44), the sequence of which is interrupted by the ‘B’ domain. The B domain appears to serve as a lid to the catalytic site, which is located on the top surface of a TIM barrel (45–51). The B domain is highly flexible (52) and adopts many conformations in the various known crystal structures. The third PYK domain (‘C’ domain) is located on the other side of the TIM barrel relative to the catalytic site and B domain. In many PYKs, the C domain can contain as many as three distinct allosteric sites (53–55). Isozymes from some species can either lack the C domain or have acquired an extra C domain (Figure 2, right) (33, 56). Isozymes similar to human liver (hLPYK) and erythrocyte (hRPYK) PYK have an extra N-terminal domain of varied length that includes a regulatory phosphorylation site, e.g. (57–59) and references therein.

Database overview

In addition to summarizing the work carried out in our laboratory on hLPYK [(6, 9, 19) and previously unpublished data], rM₁PYK (47–49, 52, 55, 60–64), ZmPYK (17) and other isozymes studied using mutagenesis, we have assembled a wide range of PYK data from the literature including sequences, structures, disease-causing substitutions, biochemical characterizations and computations (6, 9, 10, 25, 26,

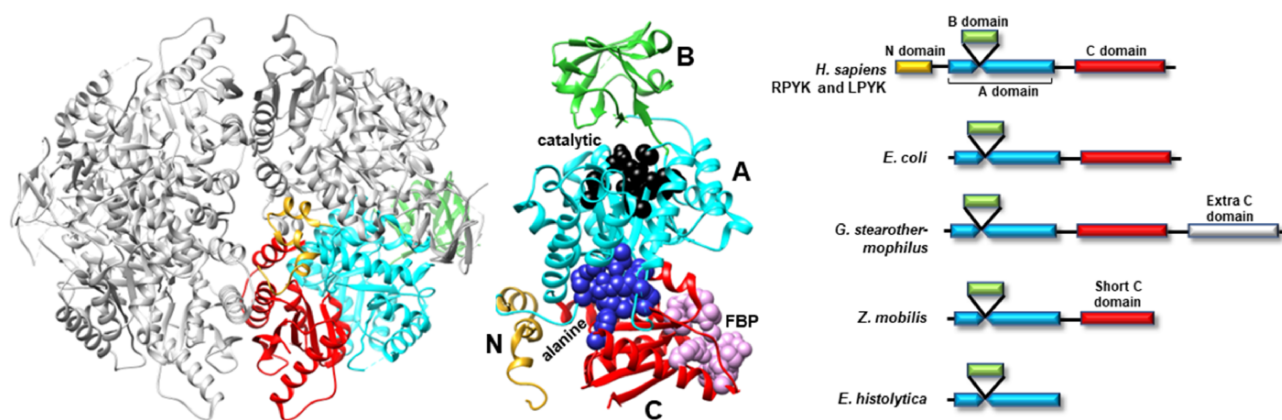


Figure 2. Representative PYK structures. Left: the homotetramer of hLPYK (PDB4IMA); three monomers are in gray, and the fourth is colored by a domain. Middle: the ribbon of the hLPYK monomer is colored by domains (A, B, C and N). The catalytic site is in black space-filling, the inhibitory alanine binding site is in blue space-filling and the enhancing Fru-1,6-BP binding site is in pink space-filling. Right: alternative domain structures observed throughout the PYK family. Domain colors match those of the middle panel.

50, 57, 58, 61, 62, 65–161). Data are presented in a format designed to facilitate their use in future efforts to advance the development of algorithms that predict substitution outcomes. The PYK-SubstitutionOME database comprises two main units, and an updateable format is maintained at <https://github.com/djparente/PYK-DB>.

Unit 1 comprises an Excel workbook, in which the first worksheet presents a summary data table and is supported by additional worksheets containing expanded data tables and links to other databases (e.g. BRENDA and GNOMAD) that contain up-to-date information on PYK. The supporting worksheets are formatted, so they can be exported as CSV files for use in various database programs. Throughout the database, data are associated with their original publication(s) using PubMed ID numbers ('PMIDs'). These PMIDs are also cited in the relevant text below; additional publications cited in the Supplemental Information include references (162–170).

Unit 2 contains three PYK sequence alignments in FASTA format. Two of these were originally published in earlier studies (9, 17, 136), and these manuscripts provide the computational details of their construction. In brief, the 'PYK Pendergrass 2006' multiple sequence alignment (MSA) contains manually curated sequences from a BLAST (173) search that sampled the whole PYK family across all domains of life; this MSA was used to compute the various bioinformatic scores reported in Unit 1 of the database. *E*-values are exceedingly low across the MSA (BLAST (171)), even though sequence identities range as low as 18%. We and others who have performed sequence analyses of this family (172–174) have found that these homologs only comprise PYK isozymes; PYK function has been verified for multiple organisms by extensive studies throughout the 1960s–1980s; many of these studies are collected from the BRENDA database [(170); <https://www.brenda-enzymes.org/enzyme.php?ecno=2.7.1.40>]; no known homologs lack the canonical PYK catalytic activity.

The second MSA ('hLPYK subfamily 2022') was generated to check that the 2006 sampling of the whole family produced comparable results (17) (i.e. the sequences deposited in 2006 provided good representation of the sequences available in 2022). Finally, using new tools developed for large datasets in

the past year, we include a new, 2000-sequence MSA for the whole family that was generated in ConSurf (176) using an HMMER search (177) of their 'Clean UniProt' database ('PYK whole family 2023 CleanUniProt.fas'). (As of this writing, UniProt contains >60,000 PYK sequences.) The comparison of ConSurf scores for the 2006 and 2023 alignments yields a Pearson correlation coefficient of 0.80 and a Spearman correlation coefficient >0.95; most of the variation occurs at the least conserved positions or the C terminus, again verifying the continued usefulness of the PYK Pendergrass 2006 MSA. (We do note that the ConSurf default value of 150 sequences provided insufficient sampling of the huge PYK family, which was determined by assessing progressively larger MSAs, until ConSurf scores converged.)

Sections of the PYK-SubstitutionOME workbook (Unit 1)

In the PYK-SubstitutionOME workbook, the first worksheet ('Features of PYK Positions') provides an overview of the whole database and summary details for each amino acid position. This main table is formatted, so that the information in Columns C and D introduces the various features (rows) and Column E through VT provide information about this feature for each PYK amino acid position. The left-most columns (A and B) contain legends to aid the interpretation of various scores. The second worksheet in this workbook, 'Legends and Definitions', contains brief descriptions of the data, definitions and expanded color legends for each of the four major sections. As further described, each row contains different types of information about the PYK positions, some of which are expanded in the supporting worksheets.

Sequences and common structural features

In the first section of the 'Features of PYK Positions' worksheet, the first set of rows contains the sequence numbers and wild-type amino acids for the main PYK isozymes that we have studied (Figure 3). The first two, hLPYK and hRPYK, are products from the same gene arising from alternative start sites. The other isozyme sequences in the main table are

Isozyme name
h=human
r=rabbit

Horizontal scroll through
positions

Position number
in native sequence

Native amino acid type

	A	B	C	D	TE	TF
1	4-Feb-23		Sequences and Common Structural Features			
2	Binding Site Key:		UNIPROT P30613-2	hLPYK Position (reference sequence)	482	483
3	PEP	Phosphoenol Pyruvate	P350A variant	hLPYK Amino Acid:	F	P
4	Ala	Alanine	UNIPROT P30613-1	hRPYK Position	513	514
5	F16BP or FBP	Fructose-1,6-Bisphosphate	P381A variant	hRPYK Amino Acid	F	P
6	Bioinformatics Key:		UNIPROT P14618-1	hM1PYK Position	470	471
7	Magenta	More Constrained Positions		hM1PYK Amino Acid	F	P
8	Green	Less Constrained Positions		rabbit M1PYK Position	470	471
9			UNIPROT P11974-1	rabbit M1PYK Amino Acid	F	P
10	Seq. Entropy:	Greater Entropy	UNIPROT O86020	ZmPYK Position	420	421
11		Lesser entropy	K189Q variant	ZmPYK Amino Acid	Q	P
12	Consurf:	Less conserved (larger score)	Map to Structure	Domain Interface	Yes	Yes
13		More conserved/higher level	From hLPYK: 4IMA	Subunit Interface		
14	TEAO (cnsvd)	Higher Score	PMID 23270483	Functional Domain	C	C
15		Lower Score		Binding Site	Ala	Ala
16	TEAO (spec)	Higher Score	Bioinformatic Scores	Seq. Entropy:	(0.00-2.82)	2.34
17		Lower Score	Whole family MSA from	Consurf Score:	(-1.15-6.44)	1.36
18	CoEv	Higher Score	PMID 16540430	Phy (TEAO - cnsvd):	(0.08-2.02)	1.05
19		Lower Score	Other Scores from	Phy (TEAO - spec):	(0.42 - 1.49)	0.74
20	EVC	Higher Score	PMIDs 34076313, 35762709	CoEvolution Score:	(0.42 - 15.37)	1.53
21		Lower Score		EVC Score (median):	(-1.62 - 2.85)	-0.20
22	2014 LP Score	Less Random (Smaller Value)		2014 Least Patterned Score (~1e-5 - ~22)	0.000	0.038
23		More Random (Bigger Value)				

Structure Features

Bioinformatics Scores
(see key at left)

Figure 3. The top two sections of the ‘Features of PYK Positions’ page from PYK-SubstitutionOME. The left-most rows contain reference information and abbreviations. The top rows (1–11) contain the position numbers and wild-type amino acids for hLPYK, hRPYK, hM1PYK, rM1PYK and ZmPYK. Note that Columns A–D are locked, allowing the position columns to be scrolled. The next section’s rows (12–23) include notes on the structural locations for each position and evolutionary scores derived from the MSAs in Unit 2 (9, 17, 136). Each evolutionary score type has a different range, which is indicated in parentheses by the score names. To aid interpretation, the scores for each position are colored to show where it falls in the range, with magenta indicating stronger constraint/more prevalent pattern and green indicating less (see the legend highlighted with the black box on the left of this figure).

human muscle PYK (hM₁PYK), rabbit muscle (rM₁PYK; a historic model system; (178, 179)) and ZmPYK. ZmPYK was included because we have studied it as an example isozyme that has different properties: ZmPYK is dimeric rather than the standard tetramer, lacks allosteric regulation and has followed an isolated evolutionary path (17). The supporting worksheet ‘Transform Homolog #s’ contains a table that aligns and numbers the analogous positions in the human, *Z. mobilis* and several other homologs contained in this database; the right side of this worksheet (yellow cells) also contains a formula to readily transform the homolog amino acid position numbering system to that of a common reference sequence (herein, hLPYK). The Python code used to create this worksheet from sequence alignments (Unit 2) is presented in Supplemental Information 2.

The next set of rows in the first section of the main worksheet associates each position with its major structural details. These include domain locations, domain and subunit contacts and direct ligand/substrate contacts. Although based on an hLPYK structure (4IMA), we expect these features to be common among most PYK isozymes. An overview of all available structures and their citations is on the supporting worksheet ‘PDB Structures Available’ (as of October 2022), along with their source organism and bound ligands (40, 44, 46–49, 53–56, 58, 60, 64, 83, 132, 139, 146, 180–216). Information about the allosteric ligands documented for other isozymes can be found at the BRENDA database (175), the link for which is included on the ‘Other Resources’ worksheet in the PYK-SubstitutionOME workbook.

Below the structural information, the third set of rows in the first section of the main table contains a range

of bioinformatic scores derived from analyses of the PYK MSA from Pendergrass 2006. The FASTA-formatted sequence alignment used to determine these scores (9, 136)—which contains additional PYK sequences not in the main table—is included as a separate text file in Unit 2. For the PYK subfamily that includes the four human isozymes, a recently expanded MSA [March 2022 (17)] is also included as a FASTA-formatted text file in Unit 2 (‘hLPYK subfamily 20022.fas’).

These different bioinformatic scores provide various ways to determine ‘non-conservation’ for each PYK position. The generation of the bioinformatic scores was previously described (9, 17). In brief, the ‘sequence entropy’ of a given position is the simplest measure of amino acid change throughout the protein family, quantifying and aggregating the frequency with which the 20 amino acids occur at that position in the alignment ($-\sum_{k=1}^{20} p_k \cdot \log p_k$, where p is the observed frequency of amino acid side-chain k at a given amino acid position) (217). ConSurf (176) scores report evolutionary rates determined from both sequence entropy and the evolutionary relatedness derived from the branch lengths of the family’s phylogenetic tree. Scores are reported on a scale of 1 (least conserved) to 9 (most conserved ‘within’ the analyzed MSA). TEAO (210) analyses are similar to those of ConSurf but reported as two sets of scores: ‘conservation’ assesses how well a position is conserved across the whole family and ‘specificity’ assesses whether a position is non-conserved in the whole family but conserved within phylogenetic branches. Co-evolutionary analyses detect pairs of positions that change in concert (219), and eigenvector centrality (‘EVC’) analyses

Horizontal scroll through positions

	A	B	C	D	TX	TY
1	4-Feb-23		Sequences and Common Structural Features			
2			UNIPROT P30613-2	hLPYK Position (reference sequence)	500	501
24			Natural Variants Identified			
25			PKLR gene	Mutation 1	C	Q
26			Disease-associated variants	Mutation 2		W
27			PMIDs on "Human disease mutants" tab	Mutation 3		P
28				Mutation 4		
29			PKM2 gene (WT seq not listed above)	Mutation 1		
30			Cancer var's PMID 24120138, 30911680	Mutation 2		
31			GNOMAD PKLR	Mutation 1	H	Q
32			All Naturally-occurring variants	Mutation 2		W
33			Downloaded Mar 25, 2020	Mutation 3		
34				Mutation 4		

Mutations in disease databases
 Mutation in natural variant database

Figure 4. The third section of the 'Features of PYK Positions' worksheet in the PYK-SubstitutionOME workbook. Rows 1 and 2 are anchored to provide position reference points, as are Columns A–D. This third section contains substitutions associated with PKD and other substitutions identified in the human population; the latter are not necessarily associated with disease.

identify positions with multiple strong partners (163). The 'least patterned' score was used to identify positions that show the least pattern of change, which has been associated with positions that have neutral substitution phenotypes (see Supplemental Information 1) (6).

Natural variants identified

Progressing down the 'Features of PYK Positions' page, the second major section notes the naturally occurring hRPYK substitutions that have been associated with disease (Figure 4). Mutations in RPYK can lead to a pyruvate kinase deficiency (PKD) that results in non-spherocytic anemia. In fact, mutations in the hRPYK protein represent the second largest set of mutations in one protein that result in medically relevant enzymopathies (220). Citations for these disease mutants are in the supporting worksheet 'Human Disease Mutants'. Although limited in number, several mutations in hM₂PKYK have been associated with cancer and are also listed in the main PYK-SubstitutionOME worksheet. The human isozymes also have other variants (not noted as disease causing) that have been observed in exome sequencing. For hRPYK, these were extracted from the GNOMAD database (221) in March 2020 for one of our studies (3) and are included in PYK-SubstitutionOME. Since the observed variants for all human isozymes are updated frequently within the GNOMAD database, we provide links to the user for the most up-to-date information (see the 'Other PYK Data Resources' supporting worksheet, which is the last worksheet of the PYK-SubstitutionOME workbook).

In vitro biochemical characterization of PYK substitutions

The third major section in 'Features of PYK Positions' summarizes the most unique data of our database: the comprehensive biochemical data of more than 1000 PYK variants that we have collected over the past decades (Figure 5). Where available, data are included for five functional parameters routinely measured in our studies (e.g. Figure 1). We note that these parameters can be affected by a wide range of conditions—buffer, pH, counter-ions, other substrates used in coupled

assays, etc. Thus, for each dataset, it is always important to compare variants to their wild-type counterpart to provide context for interpreting the change that accompanies any substitution.

Within this third section of the 'Features of PYK Positions' worksheet, the first set of rows summarizes the parameters measured in a whole-protein alanine-scanning substitution study of hLPYK (19). For reference, wild-type K and Q values measured under the same experimental conditions are included in Column D. To easily identify positions in the alanine-scanning data that alter any functional parameter, the 'Max fold-change' row reports the largest fold-change (i.e. the variant parameter divided by the equivalent wild-type parameter) observed among the five parameters; because this calculation was first developed for a study of neutral positions (6), fold-change enhancing and diminishing are reported the same (e.g. both 10-fold enhancing and 10-fold diminishing are reported as '10'). A few alanine variants lacked any enzymatic activity; these are denoted with '100' in the 'Max fold-change' row, whereas columns corresponding to unsubstituted positions (alanine or glycine in the wild-type protein) are blank.

We previously hypothesized that results from the alanine-scanning mutagenesis outcomes might be used as a proxy to identify rheostat positions: if the alanine substitution caused an outcome that was intermediate between that of wild-type and zero function, we reasoned that the position is likely to be a rheostat position (3). However, since the overall, observed range of change differed for each of the five parameters, we could use not a simple fold-change calculation to identify rheostat positions. For example, a 3-fold change in a Q parameter would be considered an intermediate value (and thus indicate a rheostat position), whereas a 3-fold change in a K parameter would not be significant.

Thus, we here normalized the alanine scan calculations, so that outcomes on the five functional parameters can be directly compared. These new results are presented on the PYK-SubstitutionOME supporting worksheet 'hLPYK Normalized AlaScan Data'; details of these new calculations are in Supplemental Information 3. From these calculations, we next determined the number of parameters (between 0 and 5) with

Horizontal scroll through positions

	A	B	C	D	UL	UM
1	4-Feb-23		Sequences and Common Structural Features			
2			UNIPROT P30613-2	hLPYK Position (reference sequence)	514	515
35			Biochemical Substitution Studies			
36			hLPYK Alanine Scan			
37			PMID: 28407397	Kix-PEP (mM) WT = 0.22	0.26	0.25
38				Kix-Ala (mM) WT = 0.25	1.2	0.30
39				Qala WT = 0.0822	0.308	0.104
40				KFBP (nM) WT = 0.11	0.28	0.17
41				QFBP WT=14.2	22.6	21.2
42				Max fold-change WT = 1 (100 = dead)	4.80	1.55
43				Number of param's in rheostat range	1	0
44	Rheostat Scale Key:	Position Classifications, Expt'l	hLPYK Maximal RheoScores (of Kpep, 2)			
45	Neutral	More neutral	PMIDs	Composite Neutral Score (0 - 1)	0.17	-
46		Less neutral	28459139	Max Weighted Rheostat Score (0 - 1)	1.00	-
47	Weighted Rheostat	More rheostatic	31740686	Max Toggle Score (0 - 1)	0.08	-
48		Less rheostatic	30117637	hLPYK Scores Above Threshold (Kpep, 1)		
49	Toggle	More toggle-like	25629396	# Neutral Scores (0.7 and above)	0	-
50		Less toggle-like	32449829	# Rheostat Scores (0.5 and above)	3	-
51			34076313	# Toggle Scores (0.64 and above)	0	-
52				ZmPYK RheoScores (Kpep)		
53			PMID	Neutral score (0 - 1)	0.31	-
54			35762709	Rheostat Score (0 - 1)	0.15	-
55				Toggle Score (0 - 1)	0.62	-
56				Assigned Phenotype	T/inactive st	-

Alanine-Scanning Mutagenesis

Toggle, Neutral, and Rheostat composite scores

Features of PYK Positions | Legends and Definitions | Transform homolog #s | PDB stru ...

Figure 5. The fourth section of the 'Features of PYK Positions' worksheet in the PYK-SubstitutionOME workbook. Again, Rows 1 and 2 and Columns A–D are anchored to provide reference points; the K and Q values for wild-type hLPYK measured under the same experimental conditions ('WT = [value]') are included for reference. This section contains information from *in vitro* biochemical characterizations of purified [and partially purified (19)] PYK variants. Both information from a whole-protein, alanine-scanning mutagenesis study and RheoScale scores from semi-SM (i.e. toggle, neutral and rheostat scores) are entered here for each position. Each RheoScale score ranges from 0 to 1; the significant ranges are indicated in parentheses, and their corresponding color legends that aid visual interpretation are noted at left. Experimentally determined parameters for >1000 substituted hLPYK proteins are included in supporting worksheets.

Horizontal scroll through positions

	A	B	C	D	RS	RT
1	4-Feb-23		Sequences and Common Structural Features			
2			UNIPROT P30613-2	hLPYK Position (reference sequence)	444	445
57			Structure-based studies			
58			Hydrogen Deuterium Exchange			
59	H/DX Key:		rM1PYK	Free Enz Rate Constant (0.0013 - 90)	90	90
60	Higher Scores:	Greater Exchange	PMID	+Phenylalanine Fold Change (4e-5 to 71.4)	4.444E-05	4.444E-05
61	Lower Scores:	Less Exchange	32479745, 23418858	+Alanine Fold Change (2e-5 to 5.4)	2.222E-05	2.222E-05
62						
63			PMID for ZmPYK: 35762709	ZmPYK #@150 min (0 to 6.14)	-	-
64			snapshot HDX data collected 2018	rM1PYK (no effectors) @180min (0 to 10.9)	10.26076	10.26076
65			Compare rM1PYK structures			
66			PDBs in Supplemental Information	Differs among Xtal structures	Y	-
67				B domain rotations		
68			Checkpoint Calculations	Summary of Q and Chi Data - FBP		
69			hLPYK	Qax	0	20.042
70			PMID	Qratio	0	1.4089
71			31010662	Chi_FBP	14164	8726
72			Experimental Q's are from alanine screen	Chi_FBP_PEP	198	0
73				Summary of Q and Chi Data - Ala		
74			PDB 4IMA	Qax	0	0.059603
75				Qratio	0	-1.3787
76				Chi_Ala	6723	21170
77				Chi_Ala_PEP	142.33	1736

H/D Exchange

Changes among M1PYK crystal structures

Checkpoint Scores

Features of PYK Positions | Legends and Definitions | Transform homolog #s | PDB stru ...

Figure 6. The bottom section of the 'Features of PYK Positions' worksheet in the PYK-SubstitutionOME workbook. Again, Rows 1 and 2 and Columns A–D are anchored to provide reference points. This section contains structural information, including H/DX data from both rM1PYK and ZmPYK, structural comparisons for M1PYK and checkpoint calculations for hLPYK. H/DX scores are color-coded to aid visual interpretation; color legends are in the left-most columns.

values that were either intermediate between wild type/dead or enhancing. This number is summarized on the 'Features of PYK Positions' worksheet in the row 'Number of

param's in rheostat range'. As observed for positions in and near the allosteric sites (10), substitutions at numerous positions in hLPYK simultaneously modulated multiple functional

parameters. Overall, >30% of all hLPYK positions appear to be rheostatic for at least one parameter.

Next, the database includes results for substitutions in hLPYK and ZmPYK. In both isozymes, many positions were assessed with semiSM mutagenesis, so that they could be assigned as having a toggle, rheostat or neutral character. As described in Supplemental Information 1, making these positional classifications requires at least 10–12 substitutions per position. Since putting all data for each variant in the main table would be unwieldy, results for individual variants are included in the supporting worksheets ‘hLPYK Variants’ and ‘ZmPYK Variants’.

To summarize the nature of each substituted position, we used the toggle, neutral and rheostat scores calculated using the RheoScore calculator (5) for each measured parameter. For non-allosteric ZmPYK, for which only K_{a-PEP} was measured, all three RheoScale scores are reported in the main worksheet for each substituted position. For allosterically regulated hLPYK, five parameters were measured for each variant and thus each position has five sets of three RheoScale scores (15 total), as well as ‘composite neutral’ scores. Composite neutral is the fraction of measured parameters that are like wild type at one position (i.e. all parameters measured for all substitutions at that position) (9). For example, for one position with 10 variants and five parameters, if 48 parameters were like wild type, the score would be $48/50 = 0.96$. In the PYK-SubstitutionOME database, the complete scores for each position are reported in the supporting worksheet ‘hLPYK RNT Scores’. Since these tables are so large, summaries—including (i) the maximal rheostat and toggle scores observed among the five parameters and the composite neutral score and (ii) the number of parameters that have scores above the phenotype threshold—as well as overall position assignments are included in the main ‘Features of PYK Positions’ worksheet (rows 45–51). The thresholds used for these assignments are described in Supplemental Information 1 and the worksheet ‘Legends and Definitions’.

The classification of position types provides one example for the utility of integrating many different characteristics of each position. For example, we previously found that rheostat positions are enriched at non-conserved positions that have an evolutionary pattern that follows speciation (9). That is, rheostat positions are conserved in subfamilies but not across the whole family of proteins. In a second example, positions with extreme ‘least patterned’ scores showed a greater propensity for neutral variants (6).

So that all hLPYK variant data can be readily searched as a single dataset, the data from the whole-protein alanine substitution and semiSM studies are also included in the ‘hLPYK Variants’ supporting worksheet, along with results for other published and unpublished hLPYK variants from our laboratory. In total, this dataset comprises more than 1000 biochemically characterized variants. The unpublished variants were generated for a variety of pilot projects and were partially purified/characterized as described by Ishwar *et al.* (25). Notably, several of the unpublished variants are associated with PKD and should be cross-referenced with the ‘Human Disease Mutants’ worksheet.

Finally, we collated a wide range of published biochemical studies performed by us and other labs for variants in PYK isozymes from a variety of species. Given their dis-

parate measurements and isozymes, we include these data in a separate supporting worksheet entitled ‘Other Isozyme Variants’. Within this section, published works are denoted by their PMIDs and each substituted isozyme position was mapped to the analogous position in hLPYK. The Python program and a description of this mapping algorithm are included in Supplemental Information 2. Note that some of these isozymes bind fructose-2,6-bisphosphate, which is designated ‘F26BP’ in the database.

Structure-based studies

In the final section of the main worksheet of the PYK-SubstitutionOME workbook (Figure 6), we report results from several studies that leverage various PYK structures. First, we include results for hydrogen/deuterium exchange as detected by mass spectrometry (H/DX-MS) for both rM₁PYK (61, 62) and ZmPYK (17). These data provide clues about which areas of the protein exchange more and less rapidly (which include information on both conformation and protein dynamics).

Next, we summarize a comparison of 61 subunits from eleven rM₁PYK structures, which were determined under a variety of liganded states and solution conditions (Figure 7; Supplemental Information 4) and can be compared to the H/DX-MS data. Side chains and backbones that showed >3.3 Å shifts were noted; the majority of the observed changes were >10 Å. Our rationale was that the differences among

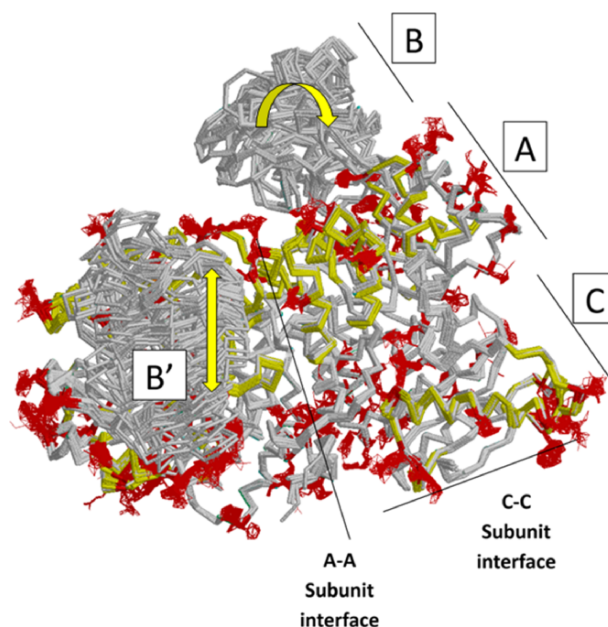


Figure 7. Locations of changes in the domains A and C were detected by comparing monomeric structures of rM1PYK. To highlight the many changes located at the A–A subunit interface, representative dimers were extracted and superimposed from each PDB; note that the perspective of the left-most monomer is looking down at the top of its B domain (B'). The other two monomers of the rM1PYK homotetramer are not shown. The locations of the largest backbone shifts are shown in light yellow (>3.3 Å; most >10 Å); and the locations of the largest side-chain differences are in dark red (>3.3 Å; most >10 Å). Since the entire B domains experience change (yellow arrows), they are not individually highlighted.

structures reflect the minimal structural ensemble, which in turn gives information about dynamics of this protein structure. It is now well recognized that protein dynamics contribute to many aspects of protein function, including catalysis, ligand binding and allosteric regulation. More details about these comparisons, along with zoomed-in images of the dynamic regions, are presented in the 'rM1PYK Structure Comparison' worksheet in the PYK-SubstitutionOME workbook.

With our rM₁PYK analysis, we also include three new rM₁PYK structures that have not previously been published: the novel liganded states are (i) a 2.4-Å structure of the M₁-PYK-Mg²⁺-Na⁺ complex, (ii) a 2.3-Å structure of the M₁-PYK-Mg²⁺-pyruvate-K⁺ complex and (iii) a 2.2-Å structure in which the various subunit active sites are populated by either a phosphonate analogue of PEP or succinate in both the presence and absence of adenosine diphosphate (ADP). Experimental details of these new structures are in Supplemental Information 5 and Supplemental Table S1. Importantly, one new structure contained a subunit bound to the phosphonate analogue of both PEP and ADP, which is the best approximation to date of the Michaelis complex for the physiologically relevant PKY reaction in any isoform.

Finally, the last rows within the 'Structure-Based Studies' section of the main PYK-SubstitutionOME worksheet summarize results of calculated mechanical coupling ('Chokepoint calculations') between active and allosteric sites of hLPYK (108). These chokepoint scores were derived by combining various adjacency matrixes determined using structural and dynamics features of hLPYK (such as Euclidean geometries and covariation in coarse-grained molecular dynamics simulations) to create a final matrix that was then analyzed for the one-site ($\chi_{\text{allosteric site}}$) and two-site ($\chi_{\text{allosteric site, PEP}}$) cost-weighted betweenness centrality parameters of the nodes. Large values indicate that a large number of pathways pass through that node, which is in turn hypothesized to identify the pathway by which allosteric communication is transmitted. The measured allosteric couplings for the whole-protein hLPYK alanine scan ('Q') were used to benchmark the chokepoint calculation.

Conclusion

The PYK-SubstitutionOME collection of biochemical characterizations for >1700 PYK substitutions, along with accompanying bioinformatic and structural information, fills a unique information niche. Indeed, very few such datasets include biochemical characterizations for multiple substitutions at many different individual protein positions (i.e. the semiSM design). In our own research, we have already used this database to draw a number of important conclusions about PYK function in particular and protein evolution in general (3, 6, 9, 10, 17, 19, 24–26, 62). We are confident that this database will be equally valuable to other researchers developing methods to predict the outcomes of amino acid substitutions. Without doubt, new measurements, methodologies and data types will be generated in future studies of PYK isozymes. We welcome input from the research community for adding such new information to the PYK-SubstitutionOME database.

Supplementary Material

Supplementary material is available at Database online.

Data availability

The PYK-SubstitutionOME database is available at <https://github.com/djparente/PYK-DB>.

Funding

W.M. Keck Foundation (to A.W.F. and L.S.K.); an internal award from the KUMC School of Medicine (to A.W.F.); the U.S. Department of Energy, Basic Energy Sciences, Office of Science (Contract No. DE-AC02-06CH11357 for use of the Advanced Photon Source); the National Center for Research Resources and the National Sciences and Engineering Research Council of Canada (to T.H.) and the National Institutes of Health (GM115340 to A.W.F., GM118589 to L.S.K. and A.W.F., DK78076 to A.W.F., P30 AG035982 to the KUMC Alzheimer's Disease Center and as a subaward to A.W.F., P20 RR17708 to T.H. and R24GM111072 for the use of BioCARS).

Conflict of interest

None declared.

Acknowledgements

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would like to thank Ms Brittany Arce, Ms Edina Kosa, Ms Brittni Jones and Ms Lisa Green (KUMC) and Dr Abby Hodges (MidAmerica Nazarene University) for assisting in populating the database and Ms Jessica Sage and Ms Laura Draxler (KUMC) for their aid in optimizing crystallization conditions.

References

- Meinhardt, S., Manley, M.W., Parente, D.J. *et al.* (2013) Rheostats and toggle switches for modulating protein function. *PLoS One*, **8**, e83502.
- Miller, M., Bromberg, Y. and Swint-Kruse, L. (2017) Computational predictors fail to identify amino acid substitution effects at rheostat positions. *Sci. Rep.*, **7**, 41329.
- Fenton, A.W., Page, B.M., Spellman-Kruse, A. *et al.* (2020) Rheostat positions: a new classification of protein positions relevant to pharmacogenomics. *Med. Chem. Res.*, **29**, 1133–1146.
- Fenton, K.D., Meneely, K.M., Wu, T. *et al.* (2022) Substitutions at a rheostat position in human aldolase A cause a shift in the conformational population. *Protein Sci.*, **31**, 357–370.
- Hodges, A.M., Fenton, A.W., Dougherty, L.L. *et al.* (2018) RheoScale: A tool to aggregate and quantify experimentally determined substitution outcomes for multiple variants at individual protein positions. *Hum. Mutat.*, **39**, 1814–1826.
- Martin, T.A., Wu, T., Tang, Q. *et al.* (2020) Identification of biochemically neutral positions in liver pyruvate kinase. *Proteins*, **88**, 1340–1350.
- Ruggiero, M.J., Malhotra, S., Fenton, A.W. *et al.* (2021) A clinically relevant polymorphism in the Na(+)/taurocholate cotransporting polypeptide (NTCP) occurs at a rheostat position. *J. Biol. Chem.*, **296**, 100047.
- Ruggiero, M.J., Malhotra, S., Fenton, A.W. *et al.* (2022) Structural plasticity is a feature of rheostat positions in the human Na(+)/taurocholate cotransporting polypeptide (NTCP). *Int. J. Mol. Sci.*, **23**, 3211.

9. Swint-Kruse, L., Martin, T.A., Page, B.M. *et al.* (2021) Rheostat functional outcomes occur when substitutions are introduced at nonconserved positions that diverge with speciation. *Protein Sci.*, **30**, 1833–1853.
10. Wu, T., Swint-Kruse, L. and Fenton, A.W. (2019) Functional tunability from a distance: rheostat positions influence allosteric coupling between two distant binding sites. *Sci. Rep.*, **9**, 16957.
11. Fowler, D.M. and Fields, S. (2014) Deep mutational scanning: a new style of protein science. *Nat. Methods*, **11**, 801–807.
12. Hietpas, R.T., Jensen, J.D. and Bolon, D.N.A. (2011) Experimental illumination of a fitness landscape. *Proc. Natl. Acad. Sci.*, **108**, 7896–7901.
13. Chan, Y.H., Venev, S.V., Zeldovich, K.B. *et al.* (2017) Correlation of fitness landscapes from three orthologous TIM barrels originates from sequence and structure constraints. *Nat. Commun.*, **8**, 14614.
14. Leander, M. *et al.* (2022) Deep mutational scanning and machine learning reveal structural and molecular rules governing allosteric hotspots in homologous proteins. *eLife*, **11**, e79932.
15. Cagiada, M., Johansson, K.E., Valanciute, A. *et al.* (2021) Understanding the origins of loss of protein function by analyzing the effects of thousands of variants on activity and abundance. *Mol. Biol. Evol.*, **38**, 3235–3246.
16. Mokhtari, D.A., Appel, M.J., Fordyce, P.M. *et al.* (2021) High throughput and quantitative enzymology in the genomic era. *Curr. Opin. Struct. Biol.*, **71**, 259–273.
17. Page, B.M., Martin, T.A., Wright, C.L. *et al.* (2022) Odd one out? Functional tuning of *Zymomonas mobilis* pyruvate kinase is narrower than its allosteric, human counterpart. *Protein Sci.*, **31**, e4336.
18. Fenton, K.D. *et al.* (2021) Substitutions at a rheostat position in human aldolase A cause a shift in the conformational population. *Protein Sci.*, **31**, 357–370.
19. Tang, Q. and Fenton, A.W. (2017) Whole-protein alanine-scanning mutagenesis of allostery: a large percentage of a protein can contribute to mechanism. *Hum. Mutat.*, **38**, 1132–1143.
20. Sousa, F.L., Parente, D.J., Shis, D.L. *et al.* (2016) AlloRep: a repository of sequence, structural and mutagenesis data for the LacI/GalR transcription regulators. *J. Mol. Biol.*, **428**, 671–678.
21. Aditham, A.K., Markin, C.J., Mokhtari, D.A. *et al.* (2021) High-throughput affinity measurements of transcription factor and DNA mutations reveal affinity and specificity determinants. *Cell Syst.*, **12**, 112–127.e11.
22. Markin, C.J., Mokhtari, D.A., Sunden, F. *et al.* (2021) Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics. *Science*, **373**, eabf8761.
23. Nisthal, A., Wang, C.Y., Ary, M.L. *et al.* (2019) Protein stability engineering insights revealed by domain-wide comprehensive mutagenesis. *Proc. Natl. Acad. Sci.*, **116**, 16367–16377.
24. Carlson, G.M. and Fenton, A.W. (2016) What mutagenesis can and cannot reveal about allostery. *Biophys. J.*, **110**, 1912–1923.
25. Ishwar, A., Tang, Q. and Fenton, A.W. (2015) Distinguishing the interactions in the fructose 1,6-bisphosphate binding site of human liver pyruvate kinase that contribute to allostery. *Biochemistry*, **54**, 1516–1524.
26. Tang, Q., Alontaga, A.Y., Holyoak, T. *et al.* (2017) Exploring the limits of the usefulness of mutagenesis in studies of allosteric mechanisms. *Hum. Mutat.*, **38**, 1144–1154.
27. Hall, E.R. and Cottam, G.L. (1978) Isozymes of pyruvate kinase in vertebrates: their physical, chemical, kinetic and immunological properties. *Int. J. Biochem.*, **9**, 785–793.
28. Ibsen, K.H. (1977) Interrelationships and functions of the pyruvate kinase isozymes and their variant forms: a review. *Cancer Res.*, **37**, 341–353.
29. Imamura, K. and Tanaka, T. (1982) Pyruvate kinase isozymes from rat. *Meth. Enzymol.*, **90 Pt E**, 150–165.
30. Blair, J. (1980) Regulatory properties of hepatic pyruvate kinase. In: Veneziale C.M. (ed.) *The Regulation of Carbohydrate Formation and Utilization in Mammals*. University Park Press, Baltimore, MD, pp. 121–151.
31. Steiner, P., Fussenegger, M., Bailey, J.E. *et al.* (1998) Cloning and expression of the *Zymomonas mobilis* pyruvate kinase gene in *Escherichia coli*. *Gene*, **220**, 31–38.
32. Pawluk, A., Scopes, R.K. and Griffiths-Smith, K. (1986) Isolation and properties of the glycolytic enzymes from *Zymomonas mobilis*. The five enzymes from glyceraldehyde-3-phosphate dehydrogenase through to pyruvate kinase. *Biochem. J.*, **238**, 275–281.
33. Kumari, P., Idrees, D., Rath, P.P. *et al.* (2020) Biochemical and biophysical characterization of the smallest pyruvate kinase from *Entamoeba histolytica*. *Biochim. Biophys. Acta - Proteins Proteom*, **1868**, 140296.
34. Fenton, A.W. (2008) Allostery: an illustrated definition for the ‘second secret of life’. *Trends Biochem. Sci.*, **33**, 420–425.
35. Reinhart, G.D. (1983) The determination of thermodynamic allosteric parameters of an enzyme undergoing steady-state turnover. *Arch. Biochem. Biophys.*, **224**, 389–401.
36. Reinhart, G.D. (1988) Linked-function origins of cooperativity in a symmetrical dimer. *Biophys. Chem.*, **30**, 159–172.
37. Reinhart, G.D. (2004) Quantitative analysis and interpretation of allosteric behavior. *Meth. Enzymol.*, **380**, 187–203.
38. Cook, P.F. and Cleland, W.W. (2007) *Enzyme Kinetics and Mechanism*. Garland Science, New York.
39. Doun, S.S., Burgner, J.W., Briggs, S.D. *et al.* (2005) *Enterococcus faecalis* phosphomevalonate kinase. *Protein Sci.*, **14**, 1134–1139.
40. Allen, S.C. and Muirhead, H. (1996) Refined three-dimensional structure of cat-muscle (M1) pyruvate kinase at a resolution of 2.6 Å. *Acta Crystallogr. D Biol. Crystallogr.*, **52**, 499–504.
41. Stammers, D.K., Levine, M., Stuart, D.I. *et al.* (1977) Structure of cat muscle pyruvate kinase at 0.26 nm resolution. *Biochem. Soc. Trans.*, **5**, 654–657.
42. Stammers, D.K. and Muirhead, H. (1975) Three-dimensional structure of cat muscle pyruvate kinase at 6 angstrom resolution. *J. Mol. Biol.*, **95**, 213–225.
43. Stammers, D.K. and Muirhead, H. (1977) Three-dimensional structure of cat muscle pyruvate kinase at 3-1 Å resolution. *J. Mol. Biol.*, **112**, 309–316.
44. Stuart, D.I., Levine, M., Muirhead, H. *et al.* (1979) Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 Å. *J. Mol. Biol.*, **134**, 109–142.
45. Oria-Hernandez, J., Cabrera, N., Pérez-Montfort, R. *et al.* (2005) Pyruvate kinase revisited: the activating effect of K⁺. *J. Biol. Chem.*, **280**, 37924–37929.
46. Morgan, H.P., McNae, I.W., Nowicki, M.W. *et al.* (2010) Allosteric mechanism of pyruvate kinase from *Leishmania mexicana* uses a rock and lock model. *J. Biol. Chem.*, **285**, 12892–12898.
47. Larsen, T.M., Benning, M.M., Rayment, I. *et al.* (1998) Structure of the Bis(Mg²⁺)-ATP-oxalate complex of the rabbit muscle pyruvate kinase at 2.1 Å resolution: ATP binding over a barrel. *Biochemistry*, **37**, 6247–6255.
48. Larsen, T.M., Benning, M.M., Wesenberg, G.E. *et al.* (1997) Ligand-induced domain movement in pyruvate kinase: structure of the enzyme from rabbit muscle with Mg²⁺, K⁺, and L-phospholactate at 2.7 Å resolution. *Arch. Biochem. Biophys.*, **345**, 199–206.
49. Larsen, T.M., Laughlin, L.T., Holden, H.M. *et al.* (1994) Structure of rabbit muscle pyruvate kinase complexed with Mn²⁺, K⁺, and pyruvate. *Biochemistry*, **33**, 6301–6309.
50. Laughlin, L.T. and Reed, G.H. (1997) The monovalent cation requirement of rabbit muscle pyruvate kinase is eliminated by substitution of lysine for glutamate 117. *Arch. Biochem. Biophys.*, **348**, 262–267.

51. Sugrue,E., Coombes,D., Wood,D. *et al.* (2020) The lid domain is important, but not essential, for catalysis of *Escherichia coli* pyruvate kinase. *Eur. Biophys. J.*, **49**, 761–772.
52. Fenton,A.W., Williams,R. and Trehwella,J. (2010) Changes in small-angle X-ray scattering parameters observed upon binding of ligand to rabbit muscle pyruvate kinase are not correlated with allosteric transitions. *Biochemistry*, **49**, 7202–7209.
53. Abdelhamid,Y., Brear,P., Greenhalgh,J. *et al.* (2019) Evolutionary plasticity in the allosteric regulator-binding site of pyruvate kinase isoform PykA from *Pseudomonas aeruginosa*. *J. Biol. Chem.*, **294**, 15505–15516.
54. Jurica,M.S., Mesecar,A., Heath,P.J. *et al.* (1998) The allosteric regulation of pyruvate kinase by fructose-1,6-bisphosphate. *Structure*, **6**, 195–210.
55. Williams,R., Holyoak,T., McDonald,G. *et al.* (2006) Differentiating a ligand's chemical requirements for allosteric interactions from those for protein binding. phenylalanine inhibition of pyruvate kinase. *Biochemistry*, **45**, 5421–5429.
56. Suzuki,K., Ito,S., Shimizu-Ibuka,A. *et al.* (2008) Crystal structure of pyruvate kinase from *Geobacillus stearothermophilus*. *J. Biochem.*, **144**, 305–312.
57. Fenton,A.W. and Tang,Q. (2009) An activating interaction between the unphosphorylated N-terminus of human liver pyruvate kinase and the main body of the protein is interrupted by phosphorylation. *Biochemistry*, **48**, 3816–3818.
58. Holyoak,T., Zhang,B., Deng,J. *et al.* (2013) Energetic coupling between an oxidizable cysteine and the phosphorylatable N-terminus of human liver pyruvate kinase. *Biochemistry*, **52**, 466–476.
59. Prasannan,C.B., Tang,Q. and Fenton,A.W. (2012) Allosteric regulation of human liver pyruvate kinase by peptides that mimic the phosphorylated/dephosphorylated N-terminus. *Methods Mol. Biol.*, **796**, 335–349.
60. Fenton,A.W., Johnson,T.A. and Holyoak,T. (2010) The pyruvate kinase model system, a cautionary tale for the use of osmolyte perturbations to support conformational equilibria in allostery. *Protein Sci.*, **19**, 1796–1800.
61. Prasannan,C.B., Gmyrek,A., Martin,T.A. *et al.* (2020) H/D exchange characterization of silent coupling: entropy-enthalpy compensation in allostery. *Biophys. J.*, **118**, 2966–2978.
62. Prasannan,C.B., Villar,M.T., Artigues,A. *et al.* (2013) Identification of regions of rabbit muscle pyruvate kinase important for allosteric regulation by phenylalanine, detected by H/D exchange mass spectrometry. *Biochemistry*, **52**, 1998–2006.
63. Schmidt-Bäse,K., Buchbinder,J.L., Reed,G.H. *et al.* (1991) Crystallization and preliminary analysis of enzyme-substrate complexes of pyruvate kinase from rabbit muscle. *Proteins*, **11**, 153–157.
64. Wooll,J.O., Friesen,R.H.E., White,M.A. *et al.* (2001) Structural and functional linkages between subunit interfaces in mammalian pyruvate kinase. *J. Mol. Biol.*, **312**, 525–540.
65. Aydın Köker,S., Oymak,Y., Bianchi,P. *et al.* (2019) A new variant of PKLR gene associated with mild hemolysis may be responsible for the misdiagnosis in pyruvate kinase deficiency. *J. Pediatr. Hematol. Oncol.*, **41**, e1–e2.
66. Bagla,S., Bhambhani,K., Gadgil,M. *et al.* (2019) Compound heterozygosity in PKLR gene for a previously unrecognized intronic polymorphism and a rare missense mutation as a novel cause of severe pyruvate kinase deficiency. *Haematologica*, **104**, e428–e431.
67. Baronciani,L. and Beutler,E. (1993) Analysis of pyruvate kinase-deficiency mutations that produce nonspherocytic hemolytic anemia. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 4324–4327.
68. Baronciani,L. and Beutler,E. (1995) Molecular study of pyruvate kinase deficient patients with hereditary nonspherocytic hemolytic anemia. *J. Clin. Invest.*, **95**, 1702–1709.
69. Baronciani,L., Bianchi,P. and Zanella,A. (1996) Hematologically important mutations: red cell pyruvate kinase (1st update). *Blood Cells Mol. Dis.*, **22**, 259–264.
70. Baronciani,L., Bianchi,P. and Zanella,A. (1996) Hematologically important mutations: red cell pyruvate kinase. *Blood Cells Mol. Dis.*, **22**, 85–89.
71. Baronciani,L., Bianchi,P. and Zanella,A. (1998) Hematologically important mutations: red cell pyruvate kinase (2nd update). *Blood Cells Mol. Dis.*, **24**, 273–279.
72. Baronciani,L., Magalhães,I.Q., Mahoney, Jr.,D.H. *et al.* (1995) Study of the molecular defects in pyruvate kinase deficient patients affected by nonspherocytic hemolytic anemia. *Blood Cells Mol. Dis.*, **21**, 49–55.
73. Berghout,J., Higgins,S., Loucoubar,C. *et al.* (2012) Genetic diversity in human erythrocyte pyruvate kinase. *Genes Immun.*, **13**, 98–102.
74. Beutler,E. and Gelbart,T. (2000) Estimating the prevalence of pyruvate kinase deficiency from the gene frequency in the general white population. *Blood*, **95**, 3585–3588.
75. Beutler,E., Westwood,B., van Zwieten,R. *et al.* (1997) G→T transition at cDNA nt 110 (K37Q) in the PKLR (pyruvate kinase) gene is the molecular basis of a case of hereditary increase of red blood cell ATP. *Hum. Mutat.*, **9**, 282–285.
76. Bianchi,P. and Fermo,E. (2020) Molecular heterogeneity of pyruvate kinase deficiency. *Haematologica*, **105**, 2218–2228.
77. Bianchi,P., Fermo,E., Lezon-Geyda,K. *et al.* (2020) Genotype-phenotype correlation and molecular heterogeneity in pyruvate kinase deficiency. *Am. J. Hematol.*, **95**, 472–482.
78. Bianchi,P. and Zanella,A. (2000) Hematologically important mutations: red cell pyruvate kinase (third update). *Blood Cells Mol. Dis.*, **26**, 47–53.
79. Bollenbach,T.J., Mesecar,A.D. and Nowak,T. (1999) Role of lysine 240 in the mechanism of yeast pyruvate kinase catalysis. *Biochemistry*, **38**, 9137–9145.
80. Bond,C.J., Jurica,M.S., Mesecar,A. *et al.* (2000) Determinants of allosteric activation of yeast pyruvate kinase and identification of novel effectors using computational screening. *Biochemistry*, **39**, 15333–15343.
81. Canu,G., De Bonis,M., Minucci,A. *et al.* (2016) Red blood cell PK deficiency: an update of PK-LR gene mutation database. *Blood Cells Mol. Dis.*, **57**, 100–109.
82. Canu,G., De Paolis,E., Righino,B. *et al.* (2020) Identification and in silico characterization of a novel PKLR genotype in a Turkish newborn. *Mol. Biol. Rep.*, **47**, 8311–8315.
83. Chen,T.J., Wang,H.-J., Liu,J.-S. *et al.* (2019) Mutations in the PKM2 exon-10 region are associated with reduced allostery and increased nuclear translocation. *Commun. Biol.*, **2**, 105.
84. Cheng,X., Friesen,R.H. and Lee,J.C. (1996) Effects of conserved residues on the regulation of rabbit muscle pyruvate kinase. *J. Biol. Chem.*, **271**, 6313–6321.
85. Cohen-Solal,M., Préhu,C., Wajcman,H. *et al.* (1998) A new sickle cell disease phenotype associating Hb S trait, severe pyruvate kinase deficiency (PK Conakry), and an alpha2 globin gene variant (Hb Conakry). *Br. J. Haematol.*, **103**, 950–956.
86. Collins,R.A., McNally,T., Fothergill-Gilmore,L.A. *et al.* (1995) A subunit interface mutant of yeast pyruvate kinase requires the allosteric activator fructose 1,6-bisphosphate for activity. *Biochem. J.*, **310** (Pt 1), 117–123.
87. Costa,C. *et al.* (2005) Severe hemolytic anemia in a Vietnamese family, associated with novel mutations in the gene encoding for pyruvate kinase. *Haematologica*, **90**, 25–30.
88. Cross,A.R., Curnutte,J.T., Rae,J. *et al.* (1996) Hematologically important mutations: X-linked chronic granulomatous disease. *Blood Cells Mol. Dis.*, **22**, 90–95.
89. Demina,A., Varughese,K.I., Barbot,J. *et al.* (1998) Six previously undescribed pyruvate kinase mutations causing enzyme deficiency. *Blood*, **92**, 647–652.

90. Ernest, I., Callens, M., Uttaro, A.D. *et al.* (1998) Pyruvate kinase of *Trypanosoma brucei*: overexpression, purification, and functional characterization of wild-type and mutated enzyme. *Protein Expr. Purif.*, **13**, 373–382.
91. Fang, M., Guo, J., Chen, D. *et al.* (2016) Halogenated carbazoles induce cardiotoxicity in developing zebrafish (*Danio rerio*) embryos. *Environ. Toxicol. Chem.*, **35**, 2523–2529.
92. Faustova, I., Loog, M. and Järvi, J. (2012) Probing L-pyruvate kinase regulatory phosphorylation site by mutagenesis. *Protein J.*, **31**, 592–597.
93. Fawaz, N., Beshlawi, I., Alqasim, A. *et al.* (2022) Novel PKLR missense mutation (A300P) causing pyruvate kinase deficiency in an Omani Kindred-PK deficiency masquerading as congenital dyserythropoietic anemia. *Clin. Case Rep.*, **10**, e05315.
94. Fenton, A.W. and Alontaga, A.Y. (2009) The impact of ions on allosteric functions in human liver pyruvate kinase. *Meth. Enzymol.*, **466**, 83–107.
95. Fenton, A.W. and Blair, J.B. (2002) Kinetic and allosteric consequences of mutations in the subunit and domain interfaces and the allosteric site of yeast pyruvate kinase. *Arch. Biochem. Biophys.*, **397**, 28–39.
96. Fermo, E., Bianchi, P., Chiarelli, L.R. *et al.* (2005) Red cell pyruvate kinase deficiency: 17 new mutations of the PK-LR gene. *Br. J. Haematol.*, **129**, 839–846.
97. Finkenstedt, A., Bianchi, P., Theurl, I. *et al.* (2009) Regulation of iron metabolism through GDF15 and hepcidin in pyruvate kinase deficiency. *Br. J. Haematol.*, **144**, 789–793.
98. Friesen, R.H., Castellani, R.J., Lee, J.C. *et al.* (1998) Allosteric in rabbit pyruvate kinase: development of a strategy to elucidate the mechanism. *Biochemistry*, **37**, 15266–15276.
99. Friesen, R.H. and Lee, J.C. (1998) The negative dominant effects of T340M mutation on mammalian pyruvate kinase. *J. Biol. Chem.*, **273**, 14772–14779.
100. Gupta, N. *et al.* (2007) Prenatal diagnosis for a novel homozygous mutation in PKLR gene in an Indian family. *Prenat. Diagn.*, **27**, 117–118.
101. Hannaert, V. *et al.* (2002) The putative effector-binding site of *Leishmania mexicana* pyruvate kinase studied by site-directed mutagenesis. *FEBS*, **514**, 255–259.
102. He, Y. *et al.* (2018) A novel PKLR gene mutation identified using advanced molecular techniques. *Pediatr. Transplant.*, **22**.
103. Ikeda, Y. and Noguchi, T. (1998) Allosteric regulation of pyruvate kinase M2 isozyme involves a cysteine residue in the intersubunit contact. *J. Biol. Chem.*, **273**, 12227–12233.
104. Ikeda, Y., Tanaka, T. and Noguchi, T. (1997) Conversion of non-allosteric pyruvate kinase isozyme into an allosteric enzyme by a single amino acid substitution. *J. Biol. Chem.*, **272**, 20495–20501.
105. Israelsen, W.J., Dayton, T., Davidson, S. *et al.* (2013) PKM2 isoform-specific deletion reveals a differential requirement for pyruvate kinase in tumor cells. *Cell*, **155**, 397–409.
106. Jamwal, M., Aggarwal, A., Das, A. *et al.* (2017) Next-generation sequencing unravels homozygous mutation in glucose-6-phosphate isomerase, GPI c.1040G > A (p.Arg347His) causing hemolysis in an Indian infant. *Clin. Chim. Acta*, **468**, 81–84.
107. Jaouani, M., Manco, L., Kalai, M. *et al.* (2017) Molecular basis of pyruvate kinase deficiency among Tunisians: description of new mutations affecting coding and noncoding regions in the PKLR gene. *Int. J. Lab. Hematol.*, **39**, 223–231.
108. Johnson, L.E., Ginovska, B., Fenton, A.W. *et al.* (2019) Choke-points in mechanical coupling associated with allosteric proteins: the pyruvate kinase example. *Biophys. J.*, **116**, 1598–1608.
109. Kager, L., Minkov, M., Zeithofer, P. *et al.* (2016) Two novel missense mutations and a 5bp deletion in the erythroid-specific promoter of the PKLR gene in two unrelated patients with pyruvate kinase deficient transfusion-dependent chronic nonspherocytic hemolytic anemia. *Pediatr. Blood Cancer*, **63**, 914–916.
110. Kanno, H., Fujii, H., Hirono, A. *et al.* (1991) cDNA cloning of human R-type pyruvate kinase and identification of a single amino acid substitution (Thr384-Met) affecting enzymatic stability in a pyruvate kinase variant (PK Tokyo) associated with hereditary hemolytic anemia. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 8218–8221.
111. Kanno, H., Fujii, H., Hirono, A. *et al.* (1992) Identical point mutations of the R-type pyruvate kinase (PK) cDNA found in unrelated PK variants associated with hereditary hemolytic anemia. *Blood*, **79**, 1347–1350.
112. Kanno, H., Fujii, H. and Miwa, S. (1993) Low substrate affinity of pyruvate kinase variant (PK Sapporo) caused by a single amino acid substitution (426 Arg→Gln) associated with hereditary hemolytic anemia. *Blood*, **81**, 2439–2441.
113. Kanno, H., Fujii, H., Tsujino, G. *et al.* (1993) Molecular basis of impaired pyruvate kinase isozyme conversion in erythroid cells: a single amino acid substitution near the active site and decreased mRNA content of the R-type PK. *Biochem. Biophys. Res. Commun.*, **192**, 46–52.
114. Kanno, H., Fujii, H., Wei, D.C. *et al.* (1997) Frame shift mutation, exon skipping, and a two-codon deletion caused by splice site mutations account for pyruvate kinase deficiency. *Blood*, **89**, 4213–4218.
115. Kanno, H., Wei, D.C., Chan, L.C. *et al.* (1994) Hereditary hemolytic anemia caused by diverse point mutations of pyruvate kinase gene found in Japan and Hong Kong. *Blood*, **84**, 3505–3509.
116. Kedar, P., Hamada, T., Warang, P. *et al.* (2009) Spectrum of novel mutations in the human PKLR gene in pyruvate kinase-deficient Indian patients with heterogeneous clinical phenotypes. *Clin. Genet.*, **75**, 157–162.
117. Kugler, W., Willaschek, C., Holtz, C. *et al.* (2000) Eight novel mutations and consequences on mRNA and protein level in pyruvate kinase-deficient patients with nonspherocytic hemolytic anemia. *Hum. Mutat.*, **15**, 261–272.
118. Lakomek, M., Huppke, P., Neubauer, B. *et al.* (1994) Mutations in the R-type pyruvate kinase gene and altered enzyme kinetic properties in patients with hemolytic anemia due to pyruvate kinase deficiency. *Ann. Hematol.*, **69**, 253–260.
119. Lenzner, C., Nuörnberg, P., Jacobasch, G. *et al.* (1997) Molecular analysis of 29 pyruvate kinase-deficient patients from Central Europe with hereditary hemolytic anemia. *Blood*, **89**, 1793–1799.
120. Lenzner, C., Nurnberg, P., Thiele, B.J. *et al.* (1994) Mutations in the pyruvate kinase L gene in patients with hereditary hemolytic anemia. *Blood*, **83**, 2817–2822.
121. Li, D.L. *et al.* (2015) A PKLR gene novel complex mutation in erythrocyte pyruvate kinase deficiency detected by targeted sequence capture and next generation sequencing. *Zhongguo Shi Yan Xue Ye Xue Za Zhi*, **23**, 1464–1468.
122. Li, H., Gu, P., Yao, R.-E. *et al.* (2014) A novel and a previously described compound heterozygous PKLR gene mutations causing pyruvate kinase deficiency in a Chinese child. *Fetal Pediatr. Pathol.*, **33**, 182–190.
123. Lin, S., Hua, X., Li, J. *et al.* (2022) Novel compound heterozygous PKLR mutation induced pyruvate kinase deficiency with persistent pulmonary hypertension in a neonate: a case report. *Front. Cardiovasc. Med.*, **9**, 872172.
124. Liu, V.M., Howell, A.J., Hosios, A.M. *et al.* (2020) Cancer-associated mutations in human pyruvate kinase M2 impair enzyme activity. *FEBS Lett.*, **594**, 646–664.
125. Lovell, S.C., Mullick, A.H. and Muirhead, H. (1998) Cooperativity in *Bacillus stearothermophilus* pyruvate kinase. *J. Mol. Biol.*, **276**, 839–851.
126. Manco, L. and Ribeiro, M.L. (2009) Novel human pathological mutations. Gene symbol: PKLR. Disease: pyruvate kinase deficiency. *Hum. Genet.*, **125**, 343.
127. Manco, L., Ribeiro, M.L., Almeida, H. *et al.* (1999) PK-LR gene mutations in pyruvate kinase deficient Portuguese patients. *Br. J. Haematol.*, **105**, 591–595.

128. Manco,L., Trovada,M.J. and Ribeiro,M.L. (2009) Novel human pathological mutations. Gene symbol: PKLR. Disease: pyruvate kinase deficiency. *Hum. Genet.*, **125**, 340.
129. Milanese,B., Pepe,C., Defelipe,L.A. *et al.* (2021) Six novel variants in the PKLR gene associated with pyruvate kinase deficiency in Argentinian patients. *Clin. Biochem.*, **91**, 26–30.
130. Mojzíkova,R., Koralkova,P., Holub,D. *et al.* (2014) Iron status in patients with pyruvate kinase deficiency: neonatal hyperferritinemia associated with a novel frameshift deletion in the PKLR gene (p.Arg518fs), and low hepcidin to ferritin ratios. *Br. J. Haematol.*, **165**, 556–563.
131. Montllor,L., Mañú-Pereira,M.D.M., Llaudet-Planas,E. *et al.* (2017) Red cell pyruvate kinase deficiency in Spain: a study of 15 cases. *Med. Clin. (Barc)*, **148**, 23–27.
132. Morgan,H.P., O'Reilly,F.J., Wear,M.A. *et al.* (2013) M2 pyruvate kinase provides a mechanism for nutrient sensing and regulation of cell proliferation. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 5881–5886.
133. Neubauer,B., Lakomek,M., Winkler,H. *et al.* (1991) Point mutations in the L-type pyruvate kinase gene of two children with hemolytic anemia caused by pyruvate kinase deficiency. *Blood*, **77**, 1871–1875.
134. Park-Hah,J.O., Kanno,H., Kim,W.D. *et al.* (2005) A novel homozygous mutation of PKLR gene in a pyruvate-kinase-deficient Korean family. *Acta Haematol.*, **113**, 208–211.
135. Pastore,L., Morte,R.D., Frisso,G. *et al.* (1998) Novel mutations and structural implications in R-type pyruvate kinase-deficient patients from Southern Italy. *Hum. Mutat.*, **11**, 127–134.
136. Pendergrass,D.C., Williams,R., Blair,J.B. *et al.* (2006) Mining for allosteric information: natural mutations and positional sequence conservation in pyruvate kinase. *IUBMB Life*, **58**, 31–38.
137. Pissard,S., Max-Audit,I., Skopinski,L. *et al.* (2006) Pyruvate kinase deficiency in France: a 3-year study reveals 27 new mutations. *Br. J. Haematol.*, **133**, 683–689.
138. Qin,L., Nie,Y., Chen,L. *et al.* (2020) Novel PLKR mutations in four families with pyruvate kinase deficiency. *Int. J. Lab. Hematol.*, **42**, e84–e87.
139. Ramírez-Silva,L., Hernández-Alcántara,G., Guerrero-Mendiola,C. *et al.* (2022) The K(+) dependent and -independent pyruvate kinases acquire the active conformation by different mechanisms. *Int. J. Mol. Sci.*, **23**, 1347.
140. Raphaël,M.F., Van Wijk,R., Schweizer,J.J. *et al.* (2007) Pyruvate kinase deficiency associated with severe liver dysfunction in the newborn. *Am. J. Hematol.*, **82**, 1025–1028.
141. Rehman,A.U., Rashid,A., Hussain,Z. *et al.* (2022) A novel homozygous missense variant p.D339N in the PKLR gene correlates with pyruvate kinase deficiency in a Pakistani family: a case report. *J. Med. Case Rep.*, **16**, 66.
142. Rouger,H., Valentin,C., Craescu,C.T. *et al.* (1996) Five unknown mutations in the LR pyruvate kinase gene associated with severe hereditary nonspherocytic haemolytic anaemia in France. *Br. J. Haematol.*, **92**, 825–830.
143. Sakai,H. (2005) Mutagenesis of the active site lysine 221 of the pyruvate kinase from *Bacillus stearothermophilus*. *J. Biochem.*, **137**, 141–145.
144. Susan-Resiga,D. and Nowak,T. (2004) Proton donor in yeast pyruvate kinase: chemical and kinetic properties of the active site Thr 298 to Cys mutant. *Biochemistry*, **43**, 15230–15245.
145. Svidnicki,M. *et al.* (2018) Novel mutations associated with pyruvate kinase deficiency in Brazil. *Rev. Bras. Hematol. Hemoter.*, **40**, 5–11.
146. Uenaka,R., Nakajima,H., Noguchi,T. *et al.* (1995) Compound heterozygous mutations affecting both hepatic and erythrocyte isozymes of pyruvate kinase. *Biochem. Biophys. Res. Commun.*, **208**, 991–998.
147. Valentini,G., Chiarelli,L., Fortin,R. *et al.* (2000) The allosteric regulation of pyruvate kinase. *J. Biol. Chem.*, **275**, 18145–18152.
148. van Solinge,W.W., Kraaijenhagen,R.J., Rijkse,G. *et al.* (1997) Molecular modelling of human red blood cell pyruvate kinase: structural implications of a novel G1091 to a mutation causing severe nonspherocytic hemolytic anemia. *Blood*, **90**, 4987–4995.
149. van Wijk,R., Huizinga,E.G. van Wesel,A.C.W. *et al.* (2009) Fifteen novel mutations in PKLR associated with pyruvate kinase (PK) deficiency: structural implications of amino acid substitutions in PK. *Hum. Mutat.*, **30**, 446–453.
150. van Wijk,R., Rijkse,G. and van Solinge,W.W. (2007) Molecular characterisation of pyruvate kinase deficiency—concerns about the description of mutant PKLR alleles. *Br. J. Haematol.*, **136**, 167–169. author reply 169–170.
151. Walker,D., Chia,W.N. and Muirhead,H. (1992) Key residues in the allosteric transition of *Bacillus stearothermophilus* pyruvate kinase identified by site-directed mutagenesis. *J. Mol. Biol.*, **228**, 265–276.
152. Zanella,A., Bianchi,P., Baronciani,L. *et al.* (1997) Molecular characterization of PK-LR gene in pyruvate kinase-deficient Italian patients. *Blood*, **89**, 3847–3852.
153. Zanella,A., Bianchi,P., Fermo,E. *et al.* (2001) Molecular characterization of the PK-LR gene in sixteen pyruvate kinase-deficient patients. *Br. J. Haematol.*, **113**, 43–48.
154. Zanella,A., Fermo,E., Bianchi,P. *et al.* (2007) Pyruvate kinase deficiency: the genotype-phenotype association. *Blood Rev.*, **21**, 217–231.
155. Zanella,A., Fermo,E., Bianchi,P. *et al.* (2005) Red cell pyruvate kinase deficiency: molecular and clinical aspects. *Br. J. Haematol.*, **130**, 11–25.
156. Zarza,R. *et al.* (1998) Molecular characterization of the PK-LR gene in pyruvate kinase deficient Spanish patients. Red Cell Pathology Group of the Spanish Society of Haematology (AEHH). *Br. J. Haematol.*, **103**, 377–382.
157. Zarza,R. *et al.* (2000) Co-existence of hereditary spherocytosis and a new red cell pyruvate kinase variant: PK Mallorca. *Haematologica*, **85**, 227–232.
158. Zarza,R. *et al.* (1999) [Molecular study of red cell pyruvate kinase deficiency in 15 patients with chronic hemolytic anemia. Group of Erythropathology of Hematology and Hemotherapy Association of Spain (HHAS)]. *Med. Clin. (Barc)*, **112**, 606–609.
159. Klei,T.R.L., Kheradmand Kia,S., Veldhuis,M. *et al.* (2017) Residual pyruvate kinase activity in PKLR-deficient erythroid precursors of a patient suffering from severe haemolytic anaemia. *Eur. J. Haematol.*, **98**, 584–589.
160. Nieto,J.M., Rochas-López,S., González-Fernández,F.A. *et al.* (2022) Next generation sequencing for diagnosis of hereditary anemia: experience in a Spanish reference center. *Clin. Chim. Acta*, **531**, 112–119.
161. Sharma,R., Jamwal,M., Senee,H.K. *et al.* (2021) Next-generation sequencing based approach to identify underlying genetic defects of Glanzmann thrombasthenia. *Indian J. Hematol. Blood Transfus.*, **37**, 414–421.
162. Larkin,M.A., Blackshields,G., Brown,N.P. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
163. Parente,D.J., Ray,J.C. and Swint-Kruse,L. (2015) Amino acid positions subject to multiple coevolutionary constraints can be robustly identified by their eigenvector network centrality scores. *Proteins*, **83**, 2293–2306.
164. Luft,J.R., Collins,R.J., Fehrman,N.A. *et al.* (2003) A deliberate approach to screening for initial crystallization conditions of biological macromolecules. *J. Struct. Biol.*, **142**, 170–179.
165. Collaborative Computational Project, Number 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.
166. Emsley,P., Lohkamp,B., Scott,W.G. *et al.* (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.

167. Chen, V.B., Arendall, W.B., Headd, J.J. *et al.* (2010) MolProbity : all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.
168. Boyer, P.D. (1962) *Pyruvate Kinase, in the Enzymes*. In: Boyer PD, Lardy HA, Myrback K eds. Academic Press, New York, pp. 95–113.
169. Urness, J.M., Clapp, K.M., Timmons, J.C. *et al.* (2013) Distinguishing the chemical moiety of phosphoenolpyruvate that contributes to allostery in muscle pyruvate kinase. *Biochemistry*, **52**, 1–3.
170. Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.*, **30**, 1022–1025.
171. Altschul, S.F. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
172. Schramm, A., Siebers, B., Tjaden, B. *et al.* (2000) Pyruvate kinase of the hyperthermophilic crenarchaeote *Thermoproteus tenax* : physiological role and phylogenetic aspects. *J. Bacteriol.*, **182**, 2001–2009.
173. Muñoz, M.E. and Ponce, E. (2003) Pyruvate kinase: current status of regulatory and functional properties. *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.*, **135**, 197–218.
174. Oria-Hernández, J., Riveros-Rosas, H. and Ramírez-Silva, L. (2006) Dichotomic phylogenetic tree of the pyruvate kinase family: K+ dependent and independent enzymes. *J. Biol. Chem.*, **281**, 30717–30724.
175. Chang, A., Jeske, L., Ulbrich, S. *et al.* (2021) BRENDA, the ELIXIR core data resource in 2021: new developments and updates. *Nucleic Acids Res.*, **49**, D498–D508.
176. Ashkenazy, H., Abadi, S., Martz, E. *et al.* (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.*, **44**, W344–W350.
177. Potter, S.C., Luciani, A., Eddy, S.R. *et al.* (2018) HMMER web server: 2018 update. *Nucleic Acids Res.*, **46**, W200–W204.
178. Boyer, P.D. (1962) *The Enzymes. Chapter 6 “Pyruvate Kinase”*, Vol. 6, 2nd edn. Academic Press, New York, pp. 95–113.
179. Kayne, F.J. (1973) 11. Pyruvate Kinase In: Boyer P.D (ed.) *The Enzymes*, Vol. 8, 3rd (8). Academic Press, New York, pp. 353–382.
180. Anastasiou, D., Yu, Y., Israelsen, W.J. *et al.* (2012) Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis. *Nat. Chem. Biol.*, **8**, 839–847.
181. Axerio-Cilies, P., See, R.H., Zoraghi, R. *et al.* (2012) Cheminformatics-driven discovery of selective, nanomolar inhibitors for staphylococcal pyruvate kinase. *ACS Chem. Biol.*, **7**, 350–359.
182. Bakszt, R., Wernimont, A., Allali-Hassani, A. *et al.* (2010) The crystal structure of *Toxoplasma gondii* pyruvate kinase 1. *PLoS One*, **5**, e12736.
183. Chaneton, B., Hillmann, P., Zheng, L. *et al.* (2012) Serine is a natural ligand and allosteric activator of pyruvate kinase M2. *Nature*, **491**, 458–462.
184. Christofk, H.R., Vander Heiden, M.G., Wu, N. *et al.* (2008) Pyruvate kinase M2 is a phosphotyrosine-binding protein. *Nature*, **452**, 181–186.
185. Cook, W.J., Senkovich, O., Aleem, K. *et al.* (2012) Crystal structure of *Cryptosporidium parvum* pyruvate kinase. *PLoS One*, **7**, e46875.
186. Dombrowska, J.D., Santarsiero, B.D. and Mesecar, A.D. (2005) Structural basis for tumor pyruvate kinase M2 allosteric regulation and catalysis. *Biochemistry*, **44**, 9417–9429.
187. Donovan, K.A., Atkinson, S.C., Kessans, S.A. *et al.* (2016) Grappling with anisotropic data, pseudo-merohedral twinning and pseudo-translational noncrystallographic symmetry: a case study involving pyruvate kinase. *Acta Crystallogr. D*, **72**, 512–519.
188. Gassaway, B.M., Cardone, R.L., Padyana, A.K. *et al.* (2019) Distinct hepatic PKA and CDK signaling pathways control activity-independent pyruvate kinase phosphorylation and hepatic glucose production. *Cell Rep.*, **29**, 3394–3404.e9.
189. Guo, C., Linton, A., Jalaie, M. *et al.* (2013) Discovery of 2-((1H-benzo[d]imidazol-1-yl)methyl)-4H-pyrido[1,2-a]pyrimidin-4-ones as novel PKM2 activators. *Bioorg. Med. Chem. Lett.*, **23**, 3358–3363.
190. Kung, C., Hixon, J., Choe, S. *et al.* (2012) Small molecule activation of PKM2 in cancer cells induces serine auxotrophy. *Chem. Biol.*, **19**, 1187–1198.
191. Matsui, Y., Yasumatsu, I., Asahi, T. *et al.* (2017) Discovery and structure-guided fragment-linking of 4-(2,3-dichlorobenzoyl)-1-methyl-pyrrole-2-carboxamide as a pyruvate kinase M2 activator. *Bioorg. Med. Chem.*, **25**, 3540–3546.
192. Mattevi, A., Valentini, G., Rizzi, M. *et al.* (1995) Crystal structure of *Escherichia coli* pyruvate kinase type I: molecular basis of the allosteric transition. *Structure*, **3**, 729–741.
193. McFarlane, J.S., Ronnebaum, T.A., Meneely, K.M. *et al.* (2019) Changes in the allosteric site of human liver pyruvate kinase upon activator binding include the breakage of an intersubunit cation- π bond. *Acta Crystallogr. F*, **75**, 461–469.
194. Morgan, H.P., McNae, I.W., Hsin, K.-Y. *et al.* (2010) An improved strategy for the crystallization of *Leishmania mexicana* pyruvate kinase. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, **66**, 215–218.
195. Morgan, H.P., McNae, I.W., Nowicki, M.W. *et al.* (2011) The trypanocidal drug suramin and other trypan blue mimetics are inhibitors of pyruvate kinases and bind to the adenosine site. *J. Biol. Chem.*, **286**, 31232–31240.
196. Morgan, H.P., Zhong, W., McNae, I.W. *et al.* (2014) Structures of pyruvate kinases display evolutionarily divergent allosteric strategies. *R. Soc. Open Sci.*, **1**, 140120.
197. Nain-Perez, A., Foller Fuchtbauer, A., Håversen, L. *et al.* (2022) Anthraquinone derivatives as ADP-competitive inhibitors of liver pyruvate kinase. *Eur. J. Med. Chem.*, **234**, 114270.
198. Nandi, S. and Dey, M. (2020) Biochemical and structural insights into how amino acids regulate pyruvate kinase muscle isoform 2. *J. Biol. Chem.*, **295**, 5390–5403.
199. Nandi, S., Razzaghi, M., Srivastava, D. *et al.* (2020) Structural basis for allosteric regulation of pyruvate kinase M2 by phosphorylation and acetylation. *J. Biol. Chem.*, **295**, 17425–17440.
200. Petchampai, N., Murillo-Solano, C., Isoe, J. *et al.* (2019) Distinctive regulatory properties of pyruvate kinase 1 from *Aedes aegypti* mosquitoes. *Insect Biochem. Mol. Biol.*, **104**, 82–90.
201. Pinto Torres, J.E., Yuan, M., Goossens, J. *et al.* (2020) Structural and kinetic characterization of *Trypanosoma congolense* pyruvate kinase. *Mol. Biochem. Parasitol.*, **236**, 111263.
202. Rigden, D.J., Phillips, S.E.V., Michels, P.A.M. *et al.* (1999) The structure of pyruvate kinase from *Leishmania mexicana* reveals details of the allosteric transition and unusual effector specificity. *J. Mol. Biol.*, **291**, 615–635.
203. Saur, M., Hartshorn, M.J., Dong, J. *et al.* (2020) Fragment-based drug discovery using cryo-EM. *Drug Discov. Today*, **25**, 485–490.
204. Schormann, N., Hayden, K.L., Lee, P. *et al.* (2019) An overview of structure, function, and regulation of pyruvate kinases. *Protein Sci.*, **28**, 1771–1784.
205. Srivastava, D., Nandi, S. and Dey, M. (2019) Mechanistic and structural insights into cysteine-mediated inhibition of pyruvate kinase muscle isoform 2. *Biochemistry*, **58**, 3669–3682.
206. Srivastava, D., Razzaghi, M., Henzl, M.T. *et al.* (2017) Structural investigation of a dimeric variant of pyruvate kinase muscle isoform 2. *Biochemistry*, **56**, 6517–6520.
207. Tulloch, L.B., Morgan, H.P., Hannaert, V. *et al.* (2008) Sulphate removal induces a major conformational change in *Leishmania mexicana* pyruvate kinase in the crystalline state. *J. Mol. Biol.*, **383**, 615–626.

208. Valentini,G., Chiarelli,L.R., Fortin,R. *et al.* (2002) Structure and function of human erythrocyte pyruvate kinase. Molecular basis of nonspherocytic hemolytic anemia. *J. Biol. Chem.*, **277**, 23807–23814.
209. Wang,P., Sun,C., Zhu,T. *et al.* (2015) Structural insight into mechanisms for dynamic regulation of PKM2. *Protein Cell*, **6**, 275–287.
210. Yuan,M., McNae,I.W., Chen,Y. *et al.* (2018) An allostatic mechanism for M2 pyruvate kinase as an amino-acid sensor. *Biochem. J.*, **475**, 1821–1837.
211. Zhong,W., Cui,L., Goh,B.C. *et al.* (2017) Allosteric pyruvate kinase-based “logic gate” synergistically senses energy and sugar levels in *Mycobacterium tuberculosis*. *Nat. Commun.*, **8**, 1986.
212. Zhong,W., Guo,J., Cui,L. *et al.* (2019) Pyruvate kinase regulates the pentose-phosphate pathway in response to hypoxia in *Mycobacterium tuberculosis*. *J. Mol. Biol.*, **431**, 3690–3705.
213. Zhong,W., Li,K., Cai,Q. *et al.* (2020) Pyruvate kinase from *Plasmodium falciparum*: structural and kinetic insights into the allosteric mechanism. *Biochem. Biophys. Res. Commun.*, **532**, 370–376.
214. Zhong,W., Morgan,H.P., McNae,I.W. *et al.* (2013) ‘In crystallo’ substrate binding triggers major domain movements and reveals magnesium as a co-activator of *Trypanosoma brucei* pyruvate kinase. *Acta Crystallogr. D Biol. Crystallogr.*, **69**, 1768–1779.
215. Zhong,W., Morgan,H., Nowicki,M. *et al.* (2014) Pyruvate kinases have an intrinsic and conserved decarboxylase activity. *Biochem. J.*, **458**, 301–311.
216. Zoraghi,R., Worrall,L., See,R.H. *et al.* (2011) Methicillin-resistant *Staphylococcus aureus* (MRSA) pyruvate kinase as a target for bis-indole alkaloids with antibacterial activities. *J. Biol. Chem.*, **286**, 44716–44725.
217. Shannon,C. (1948) The mathematical theory of communication. *Bell Syst. Tech. J.*, **27**, 379–656.
218. Ye,K., Vriend,G. and Ijzerman,A.P. (2008) Tracing evolutionary pressure. *Bioinformatics*, **24**, 908–915.
219. Parente,D.J., Swint-Kruse,L. and Gill,A.C. (2013) Multiple co-evolutionary networks are supported by the common tertiary scaffold of the LacI/GalR proteins. *PLoS One*, **8**, e84398.
220. Morado,M., Villegas,A.M., de la Iglesia,S. *et al.* (2021) Consensus document for the diagnosis and treatment of pyruvate kinase deficiency. *Med. Clin. (Barc)*, **157**, 253.e1–253.e8.
221. Karczewski,K.J. *et al.* (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *bioRxiv*, 531210.
222. Weber,G. (1972) Ligand binding and internal equilibria in proteins. *Biochemistry*, **11**, 864–878.