J Mol Model (2012) 18:2567–2576
DOI 10.1007/s00894-011-1281-3

ORIGINAL PAPER

Comparative homology modeling of pyruvate dehydrogenase kinase isozymes from *Xenopus tropicalis* reveals structural basis for their subfunctionalization

Alexander A. Tokmakov

Received: 28 July 2011 / Accepted: 13 October 2011 / Published online: 9 November 2011

Abstract Structural–functional divergence is responsible for the preservation of highly homologous genes. Protein functions affected by mutagenesis in divergent sequences require investigation on an individual basis. In the present study, comparative homology modeling and predictive bioinformatics analysis were used to reveal for the first time the subfunctionalization of two pyruvate dehydrogenase kinase (PDK) isozymes in the western clawed frog *Xenopus tropicalis*. Three-dimensional structures of the two proteins were built by homology modeling based on the crystal structures of mammalian PDKs. A detailed comparison of them revealed important structural differences that modify the accessibility of the nucleotide binding site in the two isozymes. Based on the generated models and bioinformatics data analysis, the differences between the two proteins in terms of kinetic parameters, metabolic regulation, and tissue distribution are predicted. The results obtained are consistent with the idea that one of the xtPDKs is the major isozyme responsible for metabolic control of PDC activity in *X. tropicalis*, whereas the other one has more specialized functions. Hence, this study provides a rationale for the existence of two closely related PDK isozymes in *X. tropicalis,* thereby enhancing our understanding of the functional evolution of PDK family genes.

Keywords Homology modeling · Pyruvate dehydrogenase kinase · Isoforms · Subfunctionalization · *Xenopus tropicalis*

Electronic supplementary material The online version of this article (doi:10.1007/s00894-011-1281-3) contains supplementary material, which is available to authorized users.

A. A. Tokmakov (✉)
Research Center for Environmental Genomics and Graduate School of Science, Kobe University,
Rokko dai 1-1, Nada,
Kobe 657-8501, Japan
e-mail: tokmak@phoenix.kobe-u.ac.jp

A. A. Tokmakov
RIKEN Systems and Structural Biology Center,
Yokohama, Japan

Introduction

Pyruvate dehydrogenase kinases (PDKs) inactivate the multimeric mitochondrial pyruvate dehydrogenase complex (PDC) by phosphorylating three seryl residues in the pyruvate decarboxylase subunit E1 [1, 2]. PDKs, together with the related α-ketoacid dehydrogenase kinase (BCK), constitute a distinct group of mitochondrial protein kinases (mPKs) that lack the characteristic sequence motifs of eukaryotic protein kinases [3–5]. Structural studies of PDKs and BCK have revealed that the kinases of this group bear two distinct functional domains [6, 7]. The C-terminal nucleotide-binding domain has a mixed α/β structure characteristic of the GHKL ATPase/kinase superfamily [8]. It contains four conserved motifs (G1, G2, G3, and N boxes) and the ATP-lid, which forms a unique ATP-binding fold. Nucleotide binding has been shown to induce ordering of the lid region [6, 7]. Significant differences have been observed in the conformation of the active center in the C-terminal nucleotide-binding domain of the apoenzyme, ADP-bound, and ATP-bound forms of human PDK2 [9]. The N-terminal domain of mPKs is dominated by a four-helix bundle topology and contains the lipoyl-binding pocket [10]. The important regulatory function of the PDK N-terminal domain was highlighted by the identification of at least three distinct binding sites for physiological and synthetic ligands [9]. PDKs and BCKs were demonstrated to form functional "head-to-tail" dimers through the direct interaction of the

nucleotide-binding domains of two protomers and their C-terminal tails [6, 7, 11–13]. Dimer binding to the inner lipoyl domain (L2) of the PDC E2 subunit is important for the catalytic function of PDKs [14–17]. The reversible association of the kinase with the L2 domains allows it to move along the surface of the PDC E2 core and to phosphorylate multiple copies of E1 molecules [15, 17, 18]. The oxidation and acetylation states of the lipoyl group covalently attached to L2 affect the affinity of the lipoyl domain for PDKs [12, 17].

Four genetically and biochemically distinct PDK family isozymes (PDK1, 2, 3, and 4) have been identified in mammals [19, 20]. The rationale for the existence of the four different mammalian PDK genes is currently unknown. One probable explanation is their subfunctionalization at the protein level. It was shown that the four PDK proteins have diverse substrate specificities, kinetic parameters, and tissue-specific distributions [11–13]. However, the presence of four members in the mammalian PDK family hinders comparative structure–functional studies of them.

The western clawed frog *Xenopus tropicalis* is an important model organism in developmental biology, and a viable model for genetics. The complete genome of this species has been sequenced and annotated [21]. Its relative simplicity allows studies of gene evolution and the mechanisms responsible for preserving divergent genes in genomes. The complete family of expressed PDK genes in the tissues of *X. tropicalis* consists of two members, xtPDK3 and xtPDK4, which share high sequence identity (68.4%; Fig. 1). Their 3D structures have not yet been resolved. Although divergent homologous genes have been suggested to specialize to perform complementary functions [22, 23], no evidence has been presented for the complementary function of these protein kinases. In the study described in the present paper, a homology modeling approach was combined with bioinformatics analysis to reveal for the first time the subfunctionalization of PDK proteins in *X. tropicalis*.

Methods

Sequence alignments

Amino acid sequences of *X. tropicalis* PDKs (xtPDKs) were retrieved from the UniProtKB/Swiss-Prot database (http://ca.expasy.org/sprot/, accession numbers A9ULF7 and Q6DFQ9), and template sequences were downloaded from the 1Y8P and 3CRL PDB entries. Sequence alignments were performed using the CLUSTALW (http://www.ch.embnet.org/software/ClustalW.html) [24] and FFAS03 servers (http://ffas.ljcrf.edu/ffas-cgi/cgi/pair_aln.pl) [25]. A PostScript alignment adorned with secondary structure elements was prepared with the webESPript2.2 (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) [26].

Homology modeling

Three-dimensional structures of two xtPDK isozymes were built by homology modeling based on the crystal structures of mammalian PDK proteins. The crystal structure of human PDK3 at a resolution of 2.63 Å was used as a template structure to generate the 3D model of xtPDK3. The coordinate file of the template was retrieved from the Protein Data Bank (PDB: 1Y8P chain A). The crystal structure of rat PDK2 at a resolution of 2.61 Å (PDB: 3CRL chain B) was used as the template structure for the 3D model of xtPDK4. Both template proteins were crystallized in the nucleotide-bound form. Homology modeling was carried out using the protein structure homology-modeling server SWISS-MODEL (http://swissmodel.expasy.org/) [27–29]. Visualization of the xtPDK structures was done with PYMOL [30] and DeepView/Swiss-Pdb Viewer [31].

Loop modeling

Extended loop regions are not resolved in the two PDB-deposited X-ray structures used as templates. In 1Y8P chain A, residues 311 to 319 are missing; in 3CRL chain B, residues 313 to 322 are not presented. They were initially generated in the Automatic Modeling Mode and were further refined with the Loop Builder of SWISS-MODEL.

Validation and analysis of the 3D models

The quality of the modeled xtPDK structures was evaluated using the PROCHECK [32] and QMEAN analyses [33, 34]. 3D protein structure superposition and calculations of the root mean square deviation (RMSD) values were performed using Swiss-PdbViewer by fitting carbon backbones of analyzed protein molecules. Molecular surface areas and contact volumes were calculated with the web-based tool Voss Volume Voxelator (http://www.molmovdb.org/cgi-bin/3v.cgi) [35]. B-factor profiles of the modeled proteins were investigated using the web-based tool for the analysis of protein flexibility FlexServ (http://mmb.pcb.ub.es/FlexServ/) [36], with Normal Mode Analysis employed. This server incorporates protocols for the coarse-grained determination of protein dynamics using different algorithms.

Protein structure accession numbers

The homology models of xtPDK3 and xtPDK4 proteins were submitted to the Protein Model Data Base (PMDB; http://mi.caspur.it/PMDB/) and assigned the identifiers PM0076368 and PM0076369, respectively.

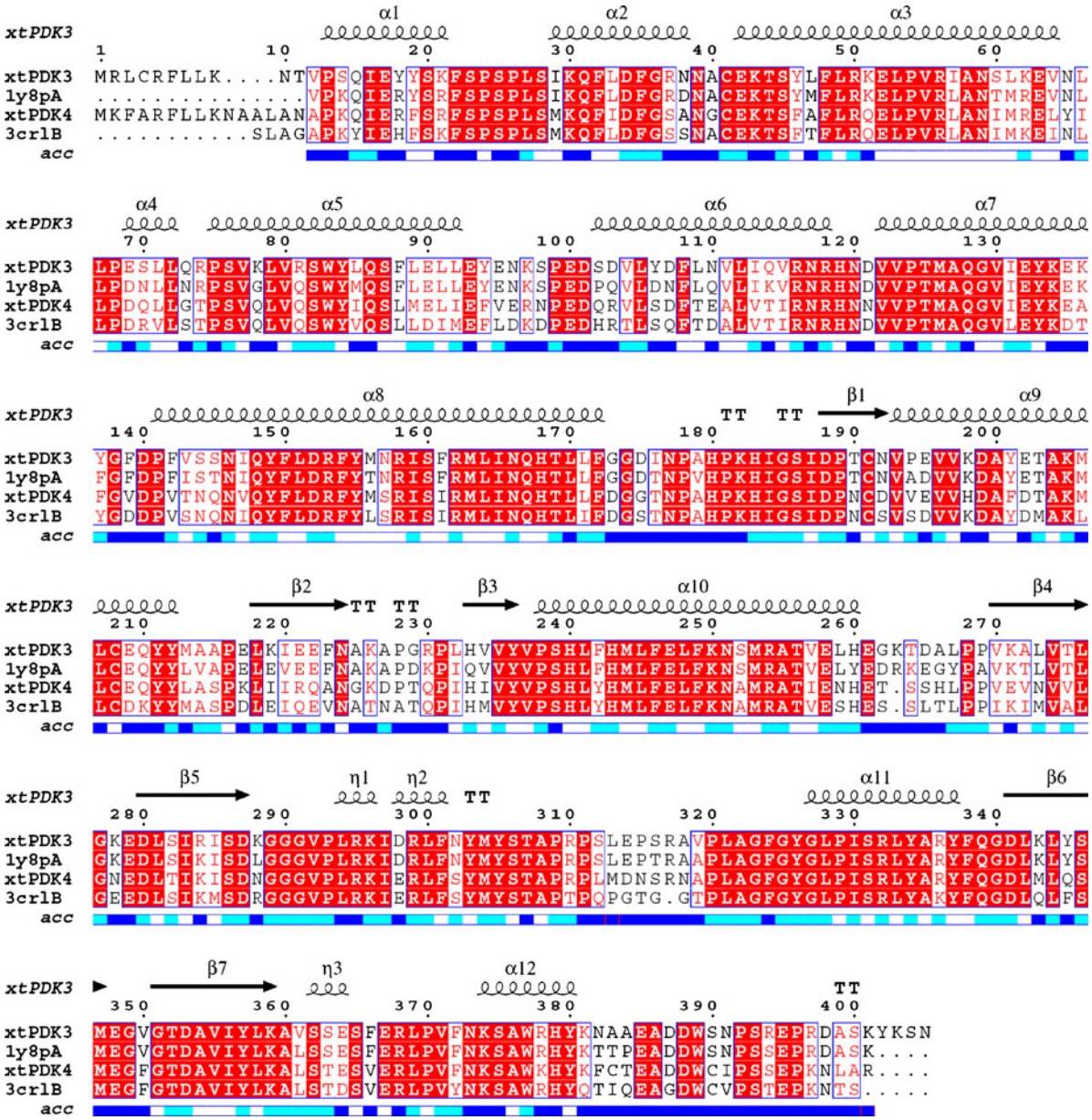


Fig. 1 Sequence alignment of *Xenopus* PDKs and their templates. Identical and similar amino acids are highlighted with a *red background and red font*, respectively. The secondary structure of xtPDK3 is annotated above the sequences. Its accessibility profile is shown below the sequences in *blue* gradations

Other methods

Energy computations were performed with the GROMOS96 implementation of Swiss-PdbViewer. Secondary structure predictions for the C-terminal regions of xtPDKs were carried out using the protein structure prediction server PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) [37, 38], the on-line proteomics server Jpred3 (http://www.compbio.dundee.ac.uk/www-jpred/) [39], and the protein disorder prediction server DISOPRED2 (http://bioinf.cs.ucl.ac.uk/disopred/) [40]. Gene expression patterns of xtPDKs inferred from EST counts were retrieved from the EST profile viewer of UniGene (http://www.ncbi.nlm.nih.gov/unigene/).

Results and discussion

Homology modeling of xtPDK3 and xtPDK4

Several X-ray structures of mammalian PDK proteins are currently deposited in the PDB. They can serve as alternative

templates for homology modeling. Template selection and target–template alignment identified the best template structures with the highest homologies and sequence coverages for model building (Tables 1, 2). The sequence identity between xtPDK3 and its template (human PDK3; PDB: 1y8pA) was 85%, whereas the identity between xtPDK4 and its template (rat PDK2; PDB: 3crlB) was 75%. These sequence identities are considered to be high enough to make reliable homology models. Sequence alignments of the two xtPDKs with their template proteins are provided in Fig. 1. The modeled 3D structures of the xtPDK3 and xtPDK4 proteins (Fig. 2) display structural features common to all PDK family enzymes, such as C-terminal and N-terminal catalytic domains, a long interdomain linker, extensive interdomain interactions, a large intermolecular cavity of the active site cleft, and an extensive loop region at the nucleotide-binding site in the C-terminal catalytic domain (the ATP lid).

The generated models of the xtPDKs were validated using PROCHECK and QMEAN analyses. Ramachandran plots of the modeled xtPDK3 and xtPDK4 structures and their statistics generated by PROCHECK are presented in the "Electronic supplementary material" (ESM) Figs. S1, S2. Calculated G factors and stereochemical parameters of the main and side chains of the models were inside the ranges of deviation in the mean values determined for the well-refined protein structures at a resolution of 2.5 Å (data not shown). Furthermore, the QMEAN scores of the models, which range between 0 and 1, with higher values indicating better quality, were calculated to be 0.805 for xtPDK3 and 0.737 for xtPDK4, indicating good overall quality of the generated models. Comparison of these scores with the normalized QMEAN scores for the available nonredundant set of PDB structures of a similar size revealed that the xtPDK models were of close-to-average quality ($|Z$ score$|<1$; Figs. S3, S4). The residues with the highest estimated errors were found to be located mainly in the loop regions of the computational models (Fig. S5). These regions were further scrutinized using coarse-grained molecular dynamics (see below). Together, both PROCHECK and QMEAN validated the folding integrities of the generated 3D models and confirmed their reliability.

Table 1 Structural templates for the homology modeling of xtPDK3. The templates with the top five scores are listed

Templates	BLAST score	Identity (%)	Coverage (%)
1Y8P (human PDK3)	715	85	98
2Q8F (human PDK1)	593	68	97
3CRL (rat PDK2)	585	67	98
2BTZ (human PDK2)	582	68	97
2ZKJ (human PDK4)	549	63	96

Table 2 Structural templates for the homology modeling of xtPDK4. The templates with the top five scores are listed

Templates	BLAST score	Identity (%)	Coverage (%)
3CRL (rat PDK2)	657	75	99
2BTZ (human PDK2)	640	75	96
2ZKJ (human PDK4)	622	73	96
2E0A (human PDK4)	618	73	96
2Q8F (human PDK1)	600	69	97

Overall comparison of the xtPDK structures

The alignment of amino acid sequences of xtPDK3 and xtPDK4 adorned with secondary structure elements is shown in Fig. 1. The sequence identity between the xtPDKs is 68.4%, which indicates quite high similarity between the two proteins. Their N-terminal regulatory domains contain eight α-helices (α1–α8), four of which (α3, α5, α6/α7, and α8) form a four-helix bundle. The C-terminal nucleotide-binding domains consist of a five-stranded β sheet (β2, and β4–7) and three α-helices (α9, α10, and α12, or α13 in xtPDK4) on one side of the β-sheet facing the N-terminal domain (Figs. 1, 3a). The N- and C-terminal domains interact with each other through a number of highly conserved hydrophobic residues in the α8 and α11 (α12 in xtPDK4) helices. These two helices are located at the bottom of a large cleft between the two domains. The long helices (α7 and α9) from each domain form the outer edges of the cleft (Fig. 3a). Since the nucleotide-binding site occupies one side of this cleft, it is considered to be the active site where the substrate binding and subsequent phosphorylation event take place [10]. The bottom part of the active site cleft is occupied by the α8/β1 loop connecting the N- and C-terminal domains.

The structures of the two xtPDK isozymes show good overall structural alignment (RMSD 1.98 Å over 368 Cα atoms), except for several mismatched loops. These include the short α2/α3 loop (residues 42–44 in xtPDK4), the N-terminal part of the interdomain linker (residues 178–182), the α10/β4 loop (residues 265–269), and a large part of the ATP lid (residues 307–327) (Fig. 3a). The misalignment of the ATP lids is especially severe. The RMSD value calculated for the superimposed xtPDK ATP lids is very

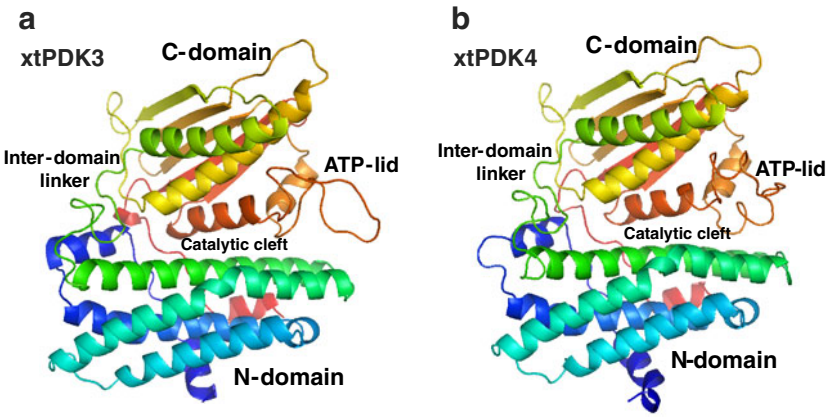


Fig. 2 **a**–**b** Modeled 3D structures of xtPDK3 (**a**) and xtPDK4 (**b**)

high (5.52 Å over 27 Cα atoms, Fig. 3b). Correspondingly, the RMSD value for the superimposed xtPDK proteins without the most misaligned regions (residues 307–323 and 326 in xtPDK4) is significantly lower than the RMSD value for the full-length proteins. It is 0.83 Å over 345 Cα atoms, indicating that the effect of the different spatial arrangements of the several mismatched loops (beyond the ATP lid) on the overall folding of the xtPDK molecule is negligible. Indeed, despite the major misalignment of the interdomain linker, which can potentially affect the relative positions of the N-terminal and C-terminal domains, the active site clefts in xtPDK3 and xtPDK4 are equally wide, as can be judged by comparing the distance between the α7 and α9 helices. Similarly, the relative arrangements of the secondary structure elements (α-helices and β-sheet strands) in the N- and C-terminal domains are not affected by misalignments in the connecting α2/α3 and α10/β4 loops.

In agreement with this analysis, an investigation of molecular flexibility using coarse-grained protein dynamics revealed that the misaligned regions of xtPDK proteins coincide mainly with the regions of high spatial uncertainty (Figs. 4a, b; S6). This approach also pinpointed the ATP lid as the major variable region in xtPDKs. In detail, the difference between the *B*-factor profiles of the ATP lids (Fig. 4c) is discussed below in the section "ATP-lid flexibility." Thus, the analysis of molecular flexibility largely supports the conclusions based on the static structure data.

In summary, despite the very high overall similarity of the modeled xtPDK structures, significant differences were observed in the folding of their ATP lids. Considering that the ATP lid is involved in nucleotide binding, the effect of the observed structural differences on this function was further evaluated.

Steric hindrance

First, we compared the spatial arrangements of the two ATP lids. Although these regions seem to be unstructured, they are stabilized by multiple intramolecular hydrogen bonds. However, the hydrogen bonding in the ATP lids of the two xtPDKs is quite different. Only six H-bonds can be created in the xtPDK3 ATP lid, whereas twelve H-bonds are formed in the ATP lid of xtPDK4 (Fig. 5). Hydrogen bonding in the ATP lid of xtPDK4 is so intense that it even

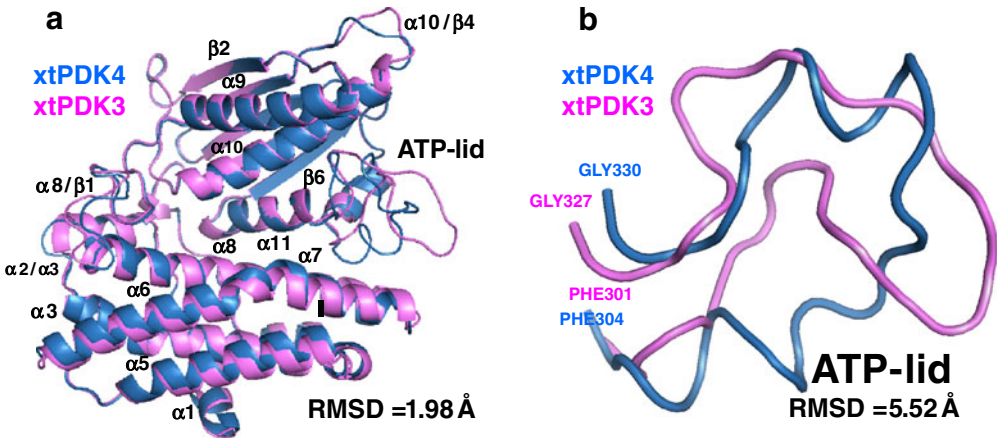
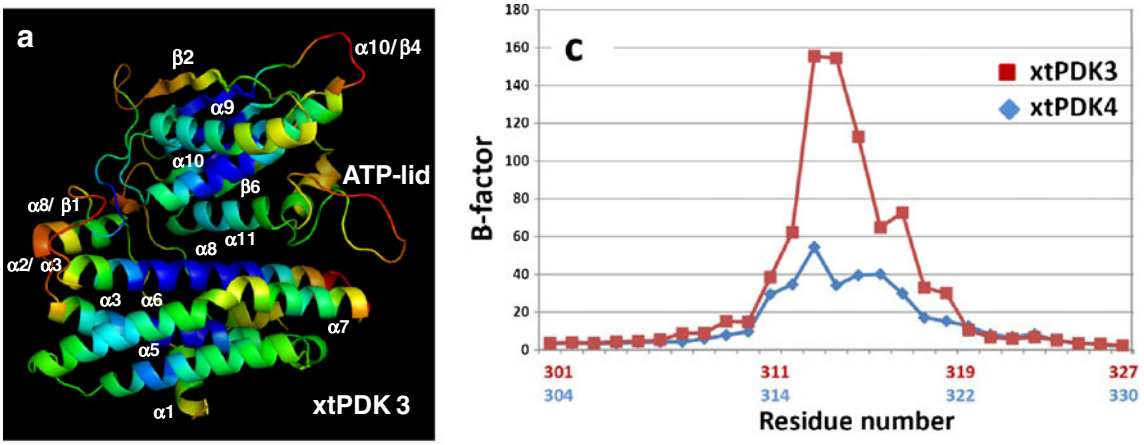


Fig. 3 **a**–**b** Superposition of the xtPDK models. Superpositions of the complete modeled regions (**a**) and the ATP lids (**b**) of the two *Xenopus* PDKs are shown



b

```
_12  V P S Q I E Y Y S K F S P S P L S I K Q F L D F G R N N A C E K T S Y L F L R K
_52  E L P V R I A N S L K E V N L L P E S L L Q R P S V K L V R S W Y L Q S F L E L
_92  L E Y E N K S P E D S D V L Y D F L N V L I Q V R N R H N D V V P T M A Q G V I
_132 E Y K E K Y G F D P F V S S N I Q Y F L D R F Y M N R I S F R M L I N Q H T L L
_172 F G G D I N P A H P K H I G S I D P T C N V P E V V K D A Y E T A K M L C E Q Y
_212 Y M A A P E L K I E E F N A K A P G R P L H V V Y V P S H L F H M L F E L F K N
_252 S M R A T V E L H E G K T D A L P P V K A L V T L G K E D L S I R I S D K G G G
_292 V P L R K I D R L F N Y M Y S T A P R P S L E P S R A V P L A G F G Y G L P I S
_332 R L Y A R Y F Q G D L K L Y S M E G V G T D A V I Y L K A V S S E S F E R L P V
_372 F N K S A W R H Y K N
```

Fig. 4 **a**–**c** Flexibility analysis of xtPDKs. The backbone and sequence of the modeled xtPDK3 protein colored by *B* factor are shown in panels **a** and **b**, respectively. *B* factors in **a** are colored based on a rainbow color spectrum, with *blue* specifying the minimum and *red* indicating the maximum. *B* factors in **b** are presented in gradations of *red*. *B*-factor landscapes of the xtPDK3 and xtPDK4 ATP lids are shown in **c**

allows the formation of two short helix-like structures in the regions of Asp317–Pro323 and Thr310–Leu315 (Fig. 3b). However, the folding of these regions is different from that of the classical α-helix, where each carbonyl oxygen is hydrogen bonded to the backbone amide, which lays four residues along the chain. The corresponding homologous

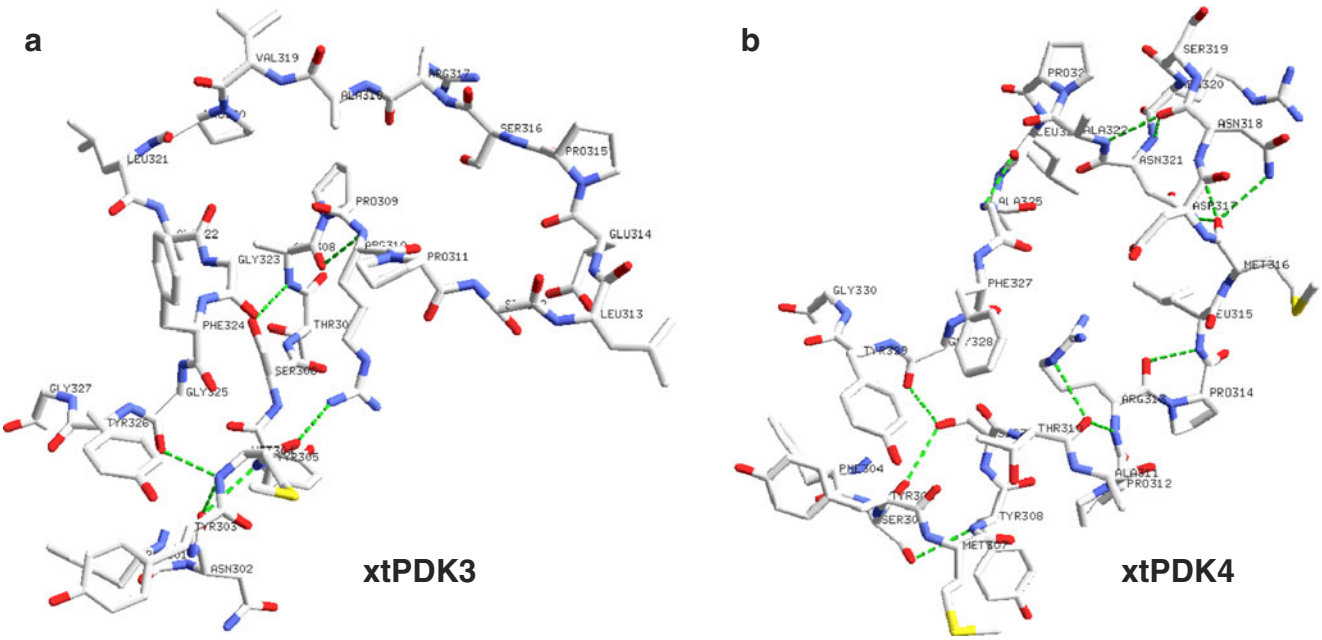


Fig. 5 **a**–**b** Hydrogen bonding in the ATP lids of xtPDK3 (**a**) and xtPDK4 (**b**). Hydrogen bonds are shown as *green dashed lines*

regions are extended in the modeled structure of xtPDK3 (Figs. 3b, 5a).

As a result of the extensive hydrogen bonding and partial structuring, the ATP lid of xtPDK4 acquires a more rigid and bulky conformation than that of xtPDK3. The calculated molecular volume of the ATP lid in xtPDK4 exceeds that in xtPDK3 (the corresponding values are 3,684 Å^3 and 3,627 Å^3). The difference (~60 Å^3) is roughly equivalent to the molecular volume of a small amino acid. Apparently, the allocation of an additional residue in the cleft between the C-terminal and N-terminal PDK domains should impose significant sterical hindrance on the nucleotide–protein interaction. These calculations suggest that the active center of xtPDK4 is less accessible to the nucleotide than that of xtPDK3.

ATP-lid flexibility

The extent of H-bonding in the ATP lid is directly related to the flexibility of this region. In turn, this may affect the accessibility of the nucleotide-binding site. In particular, the intense H-bonding in the ATP lid of the nucleotide-bound xtPDK4 (Fig. 5b) may decrease the flexibility of this region and hinder nucleotide exchange at the catalytic center. An investigation of molecular flexibility using coarse-grained protein dynamics [36] identified the ATP lid as the major variable region in xtPDKs (Figs. 4a, b; S6). A graphical representation of the *B* factor, which reflects spatial uncertainty—calculated using the web-based tool for the analysis of protein flexibility, FlexServ—for different residues of the xtPDK ATP lids is shown in Fig. 4c. The high-mobility regions of the ATP lids cover Pro311–Val319 in xtPDK3 and Pro314–Ala322 in xtPDK4, as judged by the high values of the *B* factor for these residues. Remarkably, *B*-factor values for xtPDK3 are much higher than those for xtPDK4, confirming that the intense H-bonding in the ATP lid of xtPDK4 does indeed diminish the flexibility of this region. It should be noted that the high mobility of the corresponding regions in the ATP lids is most probably related to the fact that they are flanked by the clusters of residues with high interaction energies (for details, see the section "Interaction energies in the ATP lids").

Electrostatic potentials

The ATP lids of both xtPDK isozymes contain one positive and two negative charges (Fig. 6). Thus, the resulting positive charge (+1) of this region generally favors the access of negatively charged nucleotide substrates to the active center. The presence of charged residues in the regions of high mobility strongly suggests their involvement in dynamic charge-mediated interactions with nucleotide during the catalytic cycle. Notably, the spatial distribution of the electrostatic potential is quite different in the two xtPDK proteins. The negative charge of Asp317 in xtPDK4 is buried deep inside the ATP lid and compensated by the flanking positive Arg313 and Arg320. On the other hand, the homologous negatively charged Glu314 reaches out of the ATP lid of xtPDK3. The presence of uncompensated negative charge at this location can hinder ATP access to the nucleotide binding site of xtPDK3. On the other hand, high uncompensated positive potential provided by Arg310 and Arg317 in the vicinity of the xtPDK3 nucleotide-binding site may facilitate ADP dissociation from the enzyme.

Interaction energies in the ATP lids

Changes in the standard free energy of the ATP lid upon its interaction with the nucleotide substrate contribute towards the overall energy balance of the protein kinase reaction. A comparison of the interaction energies in the ATP lids of xtPDKs can provide additional thermodynamic evidence for the subfunctionalization of these proteins. The results of interaction energy calculations are presented in the ESM Tables S1 and S2. Notably, the total interaction energies of

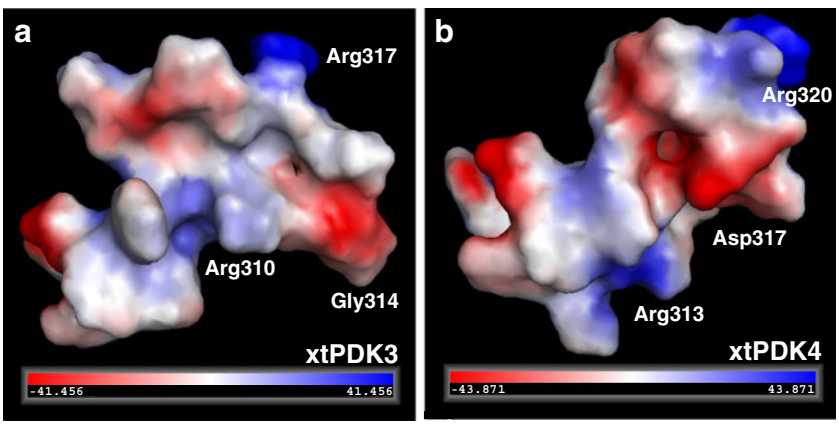


Fig. 6 **a**–**b** Electrostatic potentials in the ATP lids of xtPDK3 (**a**) and xtPDK4 (**b**). Positive potentials are shown in *blue*, negative potentials in *red*, and neutral in *white*

the high-mobility regions (Pro311–Val319 in xtPDK3 and Pro314–Ala322 in xtPDK4) differ significantly and constitute −155 kJ mol^{-1} and −226 kJ mol^{-1}, respectively. These values suggest different reaction capacities for the two ATP lids. Evidently, the high-mobility region of xtPDK4, which has a lower energy state, assumes a thermodynamically more favorable and stable conformation than that of xtPDK3. This may hinder nucleotide exchange at its catalytic center, as suggested above by the analysis of loop flexibility. Notably, the high-mobility regions in the ATP lids of both xtPDKs are flanked by clusters of residues with high interaction energies. This provides a plausible explanation for the high spatial uncertainty of the interconnecting regions. Residues in the high-energy clusters, which include Ala308, Pro309, Val319, Pro320 in xtPDK3 and Ala311, Pro312, Ala 322, Pro323, Leu324 in xtPDK4 (Tables S1, S2), will try to assume alternative conformations in order to release deposited interaction energies, thus leading to spatial rearrangements of the interconnecting loops.

Structure prediction for the C-terminal regions

The C-terminal regions of the xtPDKs (residues 381–405 in xtPDK3 and residues 384–404 in xtPDK4) were excluded from homology modeling because they are largely unstructured and partially missing from the template PDK structures deposited in the PDB. In the resolved dimer structures of several PDK family proteins, these regions represent long stretches of unfolded sequence protruding towards the opposite monomer of a dimer. The long C-terminal tail of one monomer hooks onto the opposite monomer, thus forming a structure known as “cross arms” [10]. This region is disordered in human PDK4 (PDB:2E0A), suggesting that it is highly mobile.

The amino acid sequences of the two xtPDKs differ considerably in their C-terminal regions (Fig. 1). Notably, the part following the last structural element of the protein molecule (α12 helix in xtPDK3) is four residues longer in xtPDK3 than the homologous region in xtPDK4. Secondary

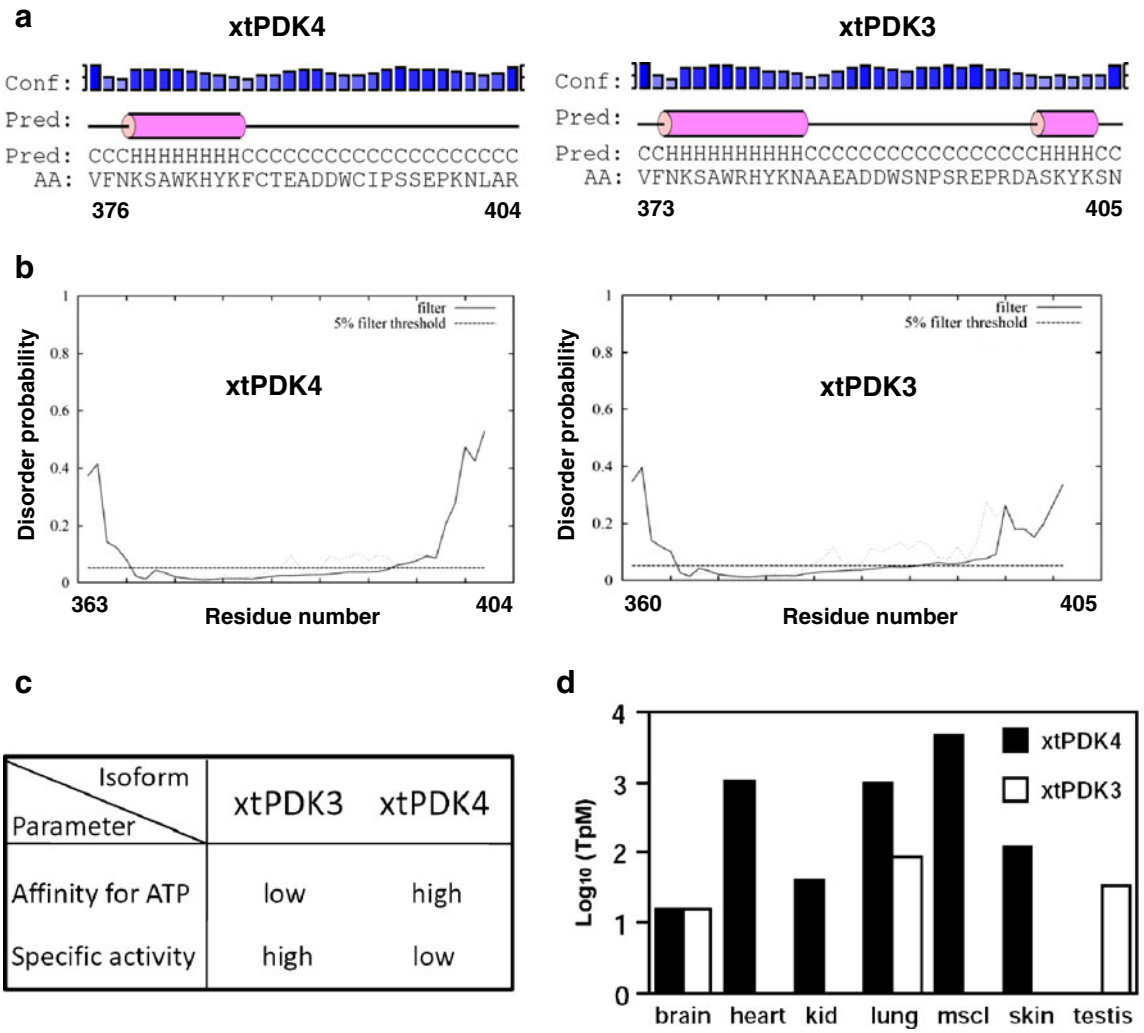


Fig. 7 **a**–**d** Bioinformatics analysis of xtPDKs. **a** Secondary structure predictions for the C-terminal regions of xtPDKs, **b** disorder predictions for the C-terminals, **c** prediction of kinetic parameters, and **d** tissue distributions of the two xtPDKs are shown

structure prediction using the protein structure prediction servers PSIPRED and Jpred3 allocates two short α-helical elements in this region of the xtPDK3 molecule, as opposed to only one α-helix in xtPDK4 (Figs. 7a, S7). Understandably, the folding of the PDK C-terminal tail, which works as a hook-and-loop fastener, affects dimer formation, and the presence of two α-helical elements in this region can be associated with stronger fastening of the xtPDK3 dimer. It should be noted, however, that although the hindmost C-terminal α-helix was predicted in xtPDK3 by both the PSIPRED and Jpred3 tools, the protein disorder prediction server DISOPRED2 assigns high disorder probability to this region in both xtPDKs (Fig. 7b). Nevertheless, in agreement with the above considerations, the greater length of the xtPDK3 C-terminal tail should promote stronger fastening of the PDK3 dimer. In turn, this may affect the recognition of the L2 lipoyl domain of PDC E2 subunit, which was shown to bind in the intermonomer cavity in reported crystal structures of PDK dimers [10, 41]. In addition, the activity of mammalian PDKs was found to be markedly enhanced through binding the lipoyl domain region [42]. Significant variation in the activation capacities of human PDK isoforms by the L2 domain has been reported [43]. Thus, it can be hypothesized that the different lengths and foldings of C-terminal tails may be associated with the different sensitivities of the xtPDKs to the L2 domain of PDC E2. As a result, the responsiveness of the two proteins to the oxidation state of this domain, which conveys the metabolic regulation of PDK activity, may differ too.

Prediction of kinetic parameters and tissue distributions of xtPDKs

The structures of xtPDK3 and xtPDK4 were modeled on the basis of the reported crystal structures of mammalian PDK3 and PDK2, respectively. The kinetic parameters of these enzymes were experimentally investigated and reported previously. The apparent K_m values for ATP for mammalian PDK2 and PDK3 are 10 μM and 50 μM, respectively, indicating that PDK2 has a much higher affinity for the nucleotide than PDK3 [11]. In addition, the specific activity of mammalian PDK2 was found to be 25-fold lower than that of PDK3 [11]. These data suggest that *Xenopus* PDKs, which are highly homologous with mammalian proteins, may also have different kinetic parameters, as shown in Fig. 7c. At present, there is no experimental evidence to support this prediction for the kinetic parameters of xtPDKs, so it should be taken with the utmost caution. However, the prediction agrees well with the performed structural analysis of the xtPDK ATP lids. Indeed, the analyses of steric hindrance, ATP-lid flexibility, electrostatic potentials, and interaction energies in the ATP lids all consistently proposed that the nucleotide substrate accessibility of xtPDK3 is greater. This should be associated with more rapid nucleotide turnover, resulting in a higher overall specific activity of xtPDK3. In addition, nucleotide binding in xtPDK3 may be of low stability and affinity, due to the structural instability of its ATP lid. Experiments with purified proteins are necessary to confirm the prediction for the kinetic parameters of xtPDKs.

An indirect confirmation of the different specific activities of the two xtPDK isozymes comes from an analysis of the tissue distributions of their mRNAs. It was previously shown that the low specific activity of mammalian PDK2 is compensated by its high expression level in various tissues [11]. On the other hand, highly active mammalian PDK3 can only be expressed at a very low level to maintain the steady-state level of PDC phosphorylation [11]. In *X. tropicalis*, the mRNA for xtPDK4 is abundant in all tissues except for testis (Fig. 7d). In contrast, xtPDK3 has a rather limited tissue distribution. Its expression level is extremely low in most tissues. Interestingly, however, xtPDK3 is predominant in testis, suggesting that it has a specialized function in this tissue. Altogether, the results of the performed analysis are consistent with the idea that xtPDK4 is the major isoenzyme responsible for metabolic control over PDC activity in *X. tropicalis,* whereas xtPDK3 must play a unique role in testes, where it is dominantly expressed.

Conclusions

The comparative homology modeling performed in this study revealed the structural basis for the subfunctionalization of the two PDK isozymes from *Xenopus tropicalis*. The study demonstrated that the two proteins have different nucleotide substrate accessibilities. In addition, a predictive bioinformatics analysis suggested that the two proteins should have different dimerization abilities, kinetic parameters, and tissue distributions. Altogether, this work discloses the structural–functional divergence of PDK proteins from *X. tropicalis* for the first time. It provides a better understanding of the divergent evolution of PDK family genes and should prove useful for structural and functional analyses of other PDK proteins.

References

1. Yeaman SJ, Hutcheson ET, Roche TE, Pettit FH, Brown JR, Reed LJ, Watson DC, Dixon GH (1978) Sites of phosphorylation on pyruvate dehydrogenase from bovine kidney and heart. Biochemistry 17:2364–2370
2. Sale GJ, Randle PJ (1981) Analysis of site occupancies in [^{32}P] phosphorylated pyruvate dehydrogenase complexes by aspartyl-prolyl cleavage of tryptic phosphopeptides. Eur J Biochem 120:535–540

3. Popov KM, Kedishvili NY, Zhao Y, Shimomura Y, Crabb DW, Harris RA (1993) Primary structure of pyruvate dehydrogenase kinase establishes a new family of eukaryotic protein kinases. J Biol Chem 268:26602–26606
4. Harris RA, Popov KM, Zhao Y, Kedishvili NY, Shimomura Y, Crabb DW (1995) A new family of protein kinases—the mitochondrial protein kinases. Adv Enzyme Regul 35:147–158
5. Bowker-Kinley M, Popov KM (1999) Evidence that pyruvate dehydrogenase kinase belongs to the ATPase/kinase superfamily. Biochem J 344:47–53
6. Machius M, Chuang JL, Wynn RM, Tomchick DR, Chuang DT (2001) Structure of rat BCKD kinase: nucleotide-induced domain communication in a mitochondrial protein kinase. Proc Natl Acad Sci USA 98:11218–11223
7. Steussy CN, Popov KM, Bowker-Kinley MM, Sloan RB Jr, Harris RA, Hamilton JA (2001) Structure of pyruvate dehydrogenase kinase. Novel folding pattern for a serine protein kinase. J Biol Chem 275:37443–37450
8. Dutta R, Inouye M (2000) GHKL, an emergent ATPase/kinase superfamily. Trends Biochem Sci 25:24–28
9. Knoechel TR, Tucker AD, Robinson CM, Phillips C, Taylor W, Bungay PJ, Kasten SA, Roche TE, Brown DG (2006) Regulatory roles of the N-terminal domain based on crystal structures of human pyruvate dehydrogenase kinase 2 containing physiological and synthetic ligands. Biochemistry 45:402–415
10. Kato M, Chuang JL, Tso SC, Wynn RM, Chuang DT (2005) Crystal structure of pyruvate dehydrogenase kinase 3 bound to lipoyl domain 2 of human pyruvate dehydrogenase complex. EMBO J 24:1763–1774
11. Bowker-Kinley MM, Davis WI, Wu P, Harris RA, Popov KM (1998) Evidence for existence of tissue-specific regulation of the mammalian pyruvate dehydrogenase complex. Biochem J 329:191–196
12. Baker JC, Yan X, Peng T, Kasten S, Roche TE (2000) Marked differences between two isoforms of human pyruvate dehydrogenase kinase. J Biol Chem 275:15773–15781
13. Korotchkina LG, Patel MS (2001) Site specificity of four pyruvate dehydrogenase kinase isoenzymes toward the three phosphorylation sites of human pyruvate dehydrogenase. J Biol Chem 276:37223–37229
14. Ono K, Radke GA, Roche TE, Rahmatullah M (1993) Partial activation of the pyruvate dehydrogenase kinase by the lipoyl domain region of E2 and interchange of the kinase between lipoyl domain regions. J Biol Chem 268:26135–26143
15. Liu S, Baker JC, Roche TE (1995) Binding of the pyruvate dehydrogenase kinase to recombinant constructs containing the inner lipoyl domain of the dihydrolipoyl acetyltransferase component. J Biol Chem 270:793–800
16. Tuganova A, Boulatnikov I, Popov KM (2002) Interaction between the individual isoenzymes of pyruvate dehydrogenase kinase and the inner lipoyl-bearing domain of transacetylase component of pyruvate dehydrogenase complex. Biochem J 366:129–136
17. Roche TE, Hiromasa Y, Turkan A, Gong X, Peng T, Yan X, Kasten SA, Bao H, Dong J (2003) Essential roles of lipoyl domains in the activated function and control of pyruvate dehydrogenase kinases and phosphatase isoform 1. Eur J Biochem 270:1050–1056
18. Hiromasa Y, Roche TE (2003) Facilitated interaction between the pyruvate dehydrogenase kinase isoform 2 and the dihydrolipoyl acetyltransferase. J Biol Chem 278:33681–33693
19. Gudi R, Bowker-Kinley MM, Kedishvili NY, Zhao Y, Popov KM (1995) Diversity of the pyruvate dehydrogenase kinase gene family in humans. J Biol Chem 270:28989–28994
20. Rowles J, Scherer SW, Xi T, Majer M, Nickle DC, Rommens JM, Popov KM, Harris RA, Riebow NL, Xia J, Tsui LC, Bogardus C, Prochazka M (1996) Cloning and characterization of PDK4 on 7q21.3 encoding a fourth pyruvate dehydrogenase kinase isoenzyme in human. J Biol Chem 271:22376–22382
21. Hellsten U et al (2010) The genome of western clawed frog *Xenopus tropicalis*. Science 328:633–636
22. Force A, Lynch M, Pickett FB, Amores A, Yan Y, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151:1531–1545
23. Lynch M, Force A (2000) The probability of duplicate gene preservation by subfunctionalization. Genetics 154:459–473
24. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acid Res 22:4673–4680
25. Jaroszewski L, Rychlewski L, Li Z, Li W, Godzik A (2005) FFAS03: a server for profile–profile sequence alignments. Nucleic Acid Res 33:W284–288
26. Gouet P, Courcelle E, Stuart D, Metoz F (1999) ESPript: Multiple sequence alignments in postscript. Bioinformatics 15:305–308
27. Peitsch MC, Wells TN, Stampf DR, Sussman JL (1995) The Swiss-3DImage collection and PDB-Browser on the World-Wide Web. Trends Biochem Sci 20:82–84
28. Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res 31:3381–3385
29. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modeling. Bioinformatics 22:195–201
30. DeLano WL (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos. http://www.pymol.org
31. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment comparative protein modeling. Electrophoresis 18:2714–2723
32. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291
33. Benkert P, Tosatto SCE, Schomburg D (2008) QMEAN: a comprehensive scoring function for model quality assessment. Proteins 71:261–277
34. Benkert P, Künzli M, Schwede T (2009) QMEAN server for protein model quality estimation. Nucleic Acids Res 37:W510–514
35. Voss NR, Gerstein M, Steitz TA, Moore PB (2006) The geometry of the ribosomal polypeptide exit tunnel. J Mol Biol 360:893–906
36. Camps J, Carrillo O, Emperador A, Orellana L, Hospital A, Rueda M, Cicin-Sain D, D'Abramo M, Gelpí JL, Orozco M (2009) FlexServ: an integrated tool for the analysis of protein flexibility. Bioinformatics 25:1709–1710
37. Bryson K, McGuffin LJ, Marsden RL, Ward JJ, Sodhi JS, Jones DT (2005) Protein structure prediction servers at University College London. Nucleic Acids Res 33:W36–38
38. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 292:195–202
39. Cole C, Barber JD, Barton GJ (2008) The Jpred 3 secondary structure prediction server server. Nucleic Acids Res 35:W197–201
40. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol 337:635–645
41. Green T, Grigoryan A, Klyuyeva A, Tuganova A, Luo M, Popov KM (2008) Structural and functional insights into the molecular mechanisms responsible for the regulation of pyruvate dehydrogenase kinase 2. J Biol Chem 283:15789–15798
42. Ono K, Radke GA, Roche TE, Rahmatullah M (1993) Partial activation of the pyruvate dehydrogenase kinase by the lipoyl domain region of E2 and interchange of the kinase between lipoyl domain regions. J Biol Chem 268:26135–26143
43. Backer JC, Yan X, Peng T, Kasten S, Roche TE (2000) Marked difference between two isoforms of human pyruvate dehydrogenase kinase. J Biol Chem 275:15773–15781