



**HAL**  
open science

# Biochemical, Bioinformatic, and Structural Comparisons of Transketolases and Position of Human Transketolase in the Enzyme Evolution

Rainier-Numa Georges, Lionel Ballut, Nushin Aghajari, Laurence Hecquet, Franck Charmantray, Bastien Doumèche

## ► To cite this version:

Rainier-Numa Georges, Lionel Ballut, Nushin Aghajari, Laurence Hecquet, Franck Charmantray, et al.. Biochemical, Bioinformatic, and Structural Comparisons of Transketolases and Position of Human Transketolase in the Enzyme Evolution. *Biochemistry*, 2024, 10.1021/acs.biochem.3c00714 . hal-04582902

**HAL Id: hal-04582902**

**<https://hal.science/hal-04582902v1>**

Submitted on 31 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Biochemical, Bioinformatic, and Structural Comparisons of Transketolases and Position of Human Transketolase in the Enzyme Evolution

Rainier-Numa Georges, Lionel Ballut, Nushin Aghajari, Laurence Hecquet, Franck Charmantray,\* and Bastien Doumèche\*



Cite This: <https://doi.org/10.1021/acs.biochem.3c00714>



Read Online

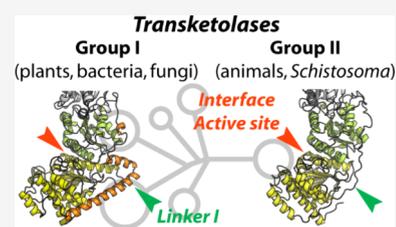
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Transketolases (TKs) are key enzymes of the pentose phosphate pathway, regulating several other critical pathways in cells. Considering their metabolic importance, TKs are expected to be conserved throughout evolution. However, Tittmann et al. (*J Biol Chem*, **2010**, 285(41): 31559–31570) demonstrated that *Homo sapiens* TK (*hsTK*) possesses several structural and kinetic differences compared to bacterial TKs. Here, we study 14 TKs from pathogenic bacteria, fungi, and parasites and compare them with *hsTK* using biochemical, bioinformatic, and structural approaches. For this purpose, six new TK structures are solved by X-ray crystallography, including the TK of *Plasmodium falciparum* (*pfTK*). All of these TKs have the same general fold as bacterial TKs. This comparative study shows that *hsTK* greatly differs from TKs from pathogens in terms of enzymatic activity, spatial positions of the active site, and monomer–monomer interface residues. An ubiquitous structural pattern is identified in all TKs as a six-residue histidyl crown around the TK cofactor (thiamine pyrophosphate), except for *hsTK* containing only five residues in the crown. Residue mapping of the monomer–monomer interface and the active site reveals that *hsTK* contains more unique residues than other TKs. From an evolutionary standpoint, TKs from animals (including *H. sapiens*) and *Schistosoma* sp. belong to a distinct structural group from TKs of bacteria, plants, fungi, and parasites, mostly based on a different linker between domains, raising hypotheses regarding evolution and regulation.



## 1. INTRODUCTION

The discovery in the 1950s of transketolases (TK, E.C. 2.2.1.1) and transaldolases (E.C. 2.2.1.2) explained how the non-oxidative part of the pentose phosphate pathway (PPP) operates.<sup>1–3</sup> PPP is ubiquitous and linked to glycolysis by three metabolites: glucose-6-phosphate (G6P), fructose-6-phosphate (F6P), and glyceraldehyde-3-phosphate (G3P). It plays a central role in the synthesis of aromatic amino acid precursors, nucleic acids, and NADPH recycling.<sup>4–6</sup> Without PPP, mitosis, metabolic and cellular functions, and resistance to oxidative stress are impossible.<sup>7–9</sup> PPP has an irreversible oxidative part and a completely reversible nonoxidative part. The reversibility of the nonoxidative part is used by pathogenic organisms and by human cancer cells to promote the neosynthesis of proteins, RNA, and DNA from F6P and G3P.<sup>10,11</sup> Furthermore, 70–80% of the PPP is mainly controlled by TK in *Homo sapiens* and *Plasmodium falciparum*. This enzyme catalyzes the reversible transfer of a two-carbon unit from a ketose (donor) to an aldose (acceptor) (Scheme 1A).<sup>1,12</sup> This unit is transiently grafted onto the enzyme's cofactor, thiamine pyrophosphate (TPP), affording  $\alpha,\beta$ -dihydroxyethylthiamine pyrophosphate (DHETPP) intermediate.

TKs are ubiquitous and appear to share many structural and enzymatic characteristics in all organisms. TKs are nested V-shaped homodimers, each monomer being formed by three

domains: the PP domain interacts with the pyrophosphate moiety of the TPP, the Pyr domain houses the thiazolium ring of the TPP in a hydrophobic pocket, and the C-ter domain promotes interactions between the monomers (see Figure 1A for a general fold of TKs).<sup>13–15</sup> TKs have two equivalent active sites, each one formed around TPP by the interaction of the PP domain of one chain with the Pyr domain of the other chain. TK PP and Pyr domains are connected by the long linker I while the short linker II connects the Pyr and C-ter domains.

In 2010, the resolution of the structure of *H. sapiens* TK (*hsTK*) by Tittmann et al. questioned the knowledge acquired on TKs.<sup>14</sup> *hsTK* has less than 30% sequence identity and a structure that differs in many aspects from *Saccharomyces cerevisiae* (*scTK*, PDB id: 1GPU), *Escherichia coli* (*ecTK*, PDB id: 2R8P), and *Mycobacterium tuberculosis* (*mtTK*, PDB id: 3R1M) TKs. The main structural difference is the absence of 62

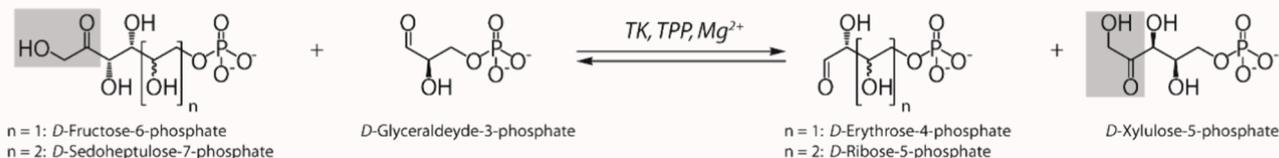
**Received:** December 19, 2023

**Revised:** May 7, 2024

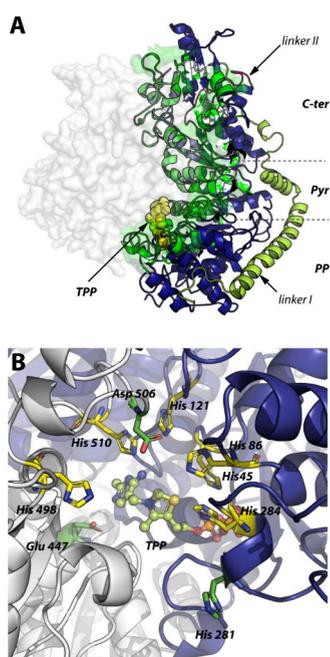
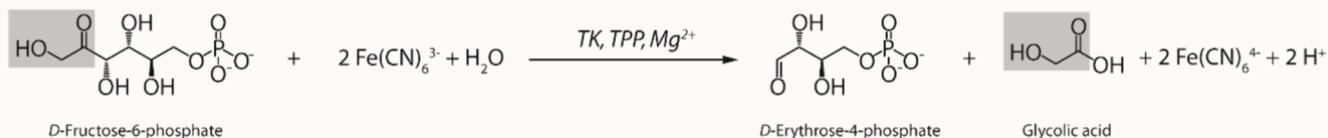
**Accepted:** May 8, 2024

**Scheme 1. (A) Reversible Reactions Catalyzed by TK in the Pentose Phosphate Pathway Underlining the Transfer of Two Carbon Atoms from a Ketose to an Aldose (Depicted as Gray Boxes); (B) Reaction of the Enzymatic Assay Showing the Release of Glycolic Acid from Fructose-6-Phosphate**

**A: Reaction catalyzed in the PPP by TK**



**B: Enzymatic assay**



**Figure 1.** (A) General fold of TKs based on the consensus structure. Monomers are depicted in blue cartoon and white transparent surface, and one TPP molecule is shown as yellow spheres. Domains of the blue monomer are indicated with dotted lines. Linker I is in limon, and linker II is in hot pink. Residues shown as green surface are residues involved in the monomer–monomer interaction as identified in Figure 2A (see Figure S4 for a back view of the enzyme). (B) TPP binding site at the interface between the two TK monomers (blue and white cartoons) showing the histidine crown (yellow sticks) and additional residues involved in TPP activation and proton stabilization (green sticks). TPP is shown as a lemon ball and stick. Nitrogen, oxygen, and phosphorus atoms are in blue, red, and orange, respectively. Residue numbers are from the TK consensus sequence (Figure S2).

63 87 residues in *hsTK* compared with *mtTK*, mainly at the end of  
64 the PP domain and in the first half of linker I, where 39  
65 residues are absent, leading to the only unstructured linker I  
66 observed up-to-date in TKs.<sup>13</sup> In contrast, linker I is formed by  
67 two (*mtTK*) or three (*ecTK* and *scTK*)  $\alpha$ -helices.<sup>13–16</sup> The  
68 other difference in *hsTK* is the absence of two loops: one in the  
69 PP domain (residues 153–167 in *mtTK*) and one in the Pyr  
70 domain (residues 413–432 in the same enzyme).

TKs follows a substituted (ping-pong) mechanism where the  
71 ketose (donor) is first converted to an aldose before the  
72 second substrate (aldose, acceptor) is converted to a ketose  
73 (Scheme 1A).<sup>16,18</sup> These transfer reactions are impossible  
74 without prior activation of the TPP into a carbanion (Figure  
75 S5). Linqvist et al. propose a model for *scTK* where TPP is  
76 activated by charge delocalization of the aminopyrimidine ring  
77 by the residue Glu 418 followed by protonation of the N1'  
78 atom of the aminopyrimidine ring.<sup>15</sup> His 481 accepts a proton  
79 from the N4' atom of the same ring, stabilizing it with the help  
80 of neighboring His 30 and His 103. The N4' atom then  
81 accepts the proton from the C2 of the thiazolium ring then  
82 forms the carbanion. More recently, another mechanism  
83 proposed by Nauton et al. involves two water molecules  
84 transferring the hydrogen atom directly from the C2 atom to  
85 the histidyl residues through a network of hydrogen bonds.<sup>19</sup>  
86 Glu 418 still initiates activation, but the N4' atom simply  
87 stabilizes the hydrogen bond network.  
88

TPP activation in *hsTK* is similar to the models presented  
89 above, except that the hydrogen bond network is extended  
90 through three water molecules, residues His 37, His 110, His  
91 258, Gln 428, and Gly 123.<sup>20</sup> The proton from the C2 of TPP  
92 is stabilized by His 37, His 110, and Gln 428, instead of His  
93 481 in *scTK*. This difference raises questions about the  
94 potentially transient position of the proton between the TPP  
95 and the His 110 that is supposed to stabilize it.  
96

From a kinetic point of view, *hsTK* has also the lowest  
97 specific activity of all known TKs: 50.4, 37.0, and 3.6  $\mu\text{mol}$   
98  $\text{min}^{-1} \text{mg}^{-1}$  for *ecTK*, *scTK*, and *hsTK*, respectively, using  
99 xylulose-5-phosphate as donor substrate and ribose-5-phos-  
100 phate as acceptor substrate.<sup>6,21–23</sup> His 481 in *scTK* is involved  
101 in the interactions with the donor substrate but not with the  
102 acceptor substrate or in the catalysis, as shown by the  
103 determination of kinetic constants. In *hsTK*, this residue  
104 position is occupied by Gln 428, which does not play a role in  
105 reaction kinetics.<sup>24</sup> *hsTK* does not accept nonphosphorylated  
106 substrates unlike other TKs such as *ecTK* or *scTK*, which  
107 accept *L*-erythrulose, for example.<sup>3,6,17</sup>  
108

Considering the differences between *hsTK* and the other  
109 described TKs, its position in evolution is questionable, being a  
110 unique enzyme or belonging to another group. This requires a  
111 more precise description of similarities and differences in  
112 structure–function relationships among TKs.  
113

This kinetic and structural comparative study is being carried out on 13 Human pathogens TKs, *sc*TK and *hs*TK. Pathogenic organisms described as priorities by the World Health Organization present a high degree of mortality risk and a very high number of cases that could lead to exchanges of genomic elements with humans. Consequently, TKs from prokaryotes (*Acinetobacter baumannii*, *Escherichia coli*, *Enterobacter faecium*, *Haemophilus influenzae*, *Mycobacterium leprae*, *Mycobacterium tuberculosis*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Streptococcus pneumoniae*) and lower and higher eukaryotes (*Candida glabrata*, *Leishmania mexicana*, *Plasmodium falciparum*) were selected together with *V. vulnificus*, a pathogenic bacterium that is lethal in 50% of cases and is currently expanding rapidly.<sup>25–30</sup> For this purpose, six new TK structures are solved by X-ray crystallography, and the others are modeled. A systematic structural study aims to identify similarities and differences in the global structure and residues forming the active sites and the monomer–monomer interfaces. For the first time, a complete mapping of these residues and sites is presented for TKs. Finally, the evolution of 134 TKs is discussed.<sup>31</sup>

## 2. EXPERIMENTAL PART

**2.1. Material.** Acrylamide/bis acrylamide 30%/0,8% (w/v), ammonium sulfate, bovine serum albumin (BSA, ≥ 98.0%), agar, dithiothreitol (DTT, ≥98.0%), EZ blue gel staining reagent, HEPES (≥99.5%), imidazole (≥99.5%), LB broth, magnesium chloride hexahydrate (≥99.0%), *N,N,N,N'*-tetramethylethylenediamine (TEMED, ≥99.0%), sodium dodecyl sulfate (SDS, ≥98.5%), sodium phosphate monobasic monohydrate (≥98.0%), thiamine chloride (≥99.0%), TPP (≥95.0%), and yeast extract (≥99.0%) are from Sigma (St Quentin Fallavier, France). Disodium hydrogen phosphate (≥99.0%), potassium hexacyanoferrate (III) ( $K_3Fe(CN)_6$ , ≥99.0%), and sodium chloride (≥99.0%) were purchased from Prolabo (Fontenay-Sous-Bois, France). Isopropyl β-D-thiogalactoside (IPTG) and PM Page ruler prestained protein ladder were from ThermoFisher (Lyon, France). Disodium D-fructose-6-phosphate hydrate (≥98.0%) was from ABCR (Lyon, France) and potassium chloride (≥97.0%) was from BDH (VWR, Rosny-sous-Bois, France). Bradford protein assay dye reagent concentrate was obtained from Bio-Rad (Marnes-la-Coquette, France). NuPage LDS loading sample buffer 4X was purchased from Invitrogen (Lyon, France). Sodium ampicillin (≥97.0%) was from Roth. TK genes were synthesized according to their UniProt sequences and inserted in plasmids by GenScript (Rijswijk, The Netherlands) (*vide infra*). Chemically competent *E. coli* BL21(DE3) cells were prepared in our lab and electro competent *E. coli* BL21(DE3) were from NEB. His trap FF 5 mL affinity column and Hi-Prep 26/10 desalting columns were purchased from GE Healthcare (Tremblay-en-France, France).

**2.2. TK Sequences Selection.** TKs protein sequences were collected from UniProt database (<https://www.uniprot.org/>) and the Protein Data Bank (PDB, <https://www.rcsb.org/>) using the keyword “Transketolase”, and the E.C. number: 2.2.1.1. 273,175 entries were obtained from UniProt and 398 from the PDB. Data curation on UniProt entries was carried out by discarding all TKs from nonpathogenic organisms for humans. All sequences whose length was not between 600 and 750 residues and therefore assumed to be truncated TK were also discarded.

Twenty-eight TKs sequences were selected, plus the *hs*TK already crystallized by Tittmann et al.<sup>14</sup> To ensure the completeness of the sequences, all TKs cDNAs sequences accessed from UniProt were translated into a protein sequence using APE v2.0.61 (<https://jorgensen.biology.utah.edu/wayned/a-pe/>) and aligned with their own protein sequence referenced in UniProt using Stretcher ([https://www.ebi.ac.uk/Tools/psa/emboss\\_stretcher/](https://www.ebi.ac.uk/Tools/psa/emboss_stretcher/); alignment matrix: BLOSUM62, gap opened: 1, gap extend: 1). Structure–function preservation was accepted if a TK sequence had at least 30% sequence identity with a yet crystallized TK: *hs*TK (PDB id 3MOS) or *ec*TK (PDB id 2R8P). Retained TKs are shown in Table 1.

**Table 1. TKs Selected in This Work with Their Abbreviation, UniProt Accession Number, and PDB Template Used for Homology Modelling *ec*TK, *hs*TK, *lm*TK, and *mt*TK Were Already Available on the PDB<sup>a</sup>**

organism	abbreviation	UniProt code	PDB template
<i>Acinetobacter baumannii</i>	<i>ab</i> TK	B0VR28	2R8P
<i>Candida glabrata</i>	<i>cg</i> TK	Q6FWC3	1TRK
<i>Escherichia coli</i>	<i>ec</i> TK	P27302	n.a.
<i>Enterococcus faecium</i>	<i>ef</i> TK	I3UIP7	4C7 V
<i>Haemophilus influenzae</i>	<i>hi</i> TK	P43757	2R8P
<i>Homo sapiens</i>	<i>hs</i> TK	P29401	n.a.
<i>Leishmania mexicana</i>	<i>lm</i> TK	Q8MMP3	n.a.
<i>Mycobacterium leprae</i>	<i>ml</i> TK	P46708	3RIM
<i>Mycobacterium tuberculosis</i>	<i>mt</i> TK	P9WG25	n.a.
<i>Plasmodium falciparum</i>	<i>pf</i> TK	C6KSV3	5HJE
<i>Pseudomonas aeruginosa</i>	<i>pa</i> TK	Q9ISY8	2R8P
<i>Staphylococcus aureus</i>	<i>sa</i> TK	P99161	3HYL
<i>Streptococcus pneumoniae</i>	<i>sp</i> TK	P22976	3HYL

<sup>a</sup>n.a.: not applicable.

**2.3. Plasmids.** The TEV protease recognition sequence was added in silico to the C-term of each TK. No cleavage sites for TEV protease were found on TKs sequences using Peptide Cutter ([https://web.expasy.org/peptide\\_cutter/](https://web.expasy.org/peptide_cutter/)) on the contrary to enterokinase, thrombin, and factor Xa, and with the exception of *P. aeruginosa* TK (*pa*TK) that contains a TEV protease cleavage site (Nter–ENLYFQG–Cter). In this case, the protease recognition sequence chosen was enterokinase (Nter–DDDKI–Cter). Corresponding cDNA sequences were optimized for expression in *E. coli*, synthesized, and inserted into plasmid pET21b(+) between NdeI 5' and XhoI 3' restriction sites by GenScript (Rijswijk, The Netherlands). This plasmid included a His<sub>6</sub>-tag coding sequence after the XhoI restriction site. The protease cleavage site was included with the intention to obtain a native protein in case of the tag interfered with crystallization. However, since uncleaved proteins allowed the acquisition of structures at a satisfactory resolution, the decision was made to retain the 6× His tag.

**2.4. Protein Expression and Purification.** Chemical competent *E. coli* BL21 (DE3) strain was incubated with 200 ng of plasmid. After 10 min on ice, bacteria were transformed by heat shock at 42 °C for 45 s. Bacteria were recovered for 1 h in LB medium at 37 °C under agitation and then centrifuged at 10,000 × *g* at room temperature for 1 min. Transformed bacteria were suspended in 100 μL of LB and spread on an LB-agar plate supplemented with 100 μg mL<sup>-1</sup> ampicillin. Clones were grown overnight at 37 °C. The electrocompetent *E. coli* BL21 (DE3) strain was used to obtain *ec*TK clones. 214

215 Precultured (LB, 100  $\mu\text{g mL}^{-1}$  ampicillin, overnight at 37 °C)  
216 were seeded with a transformed clone. Then, 2 L cultures were  
217 carried out in LB at 30 °C under agitation supplemented with  
218 100  $\mu\text{g mL}^{-1}$  ampicillin from a 1/100 (v/v) preculture.  
219 Induction was done once the OD<sup>600 nm</sup> reached 0.6–0.8 by the  
220 addition of IPTG (0.5 mM) and thiamine (10  $\mu\text{M}$ ).  
221 Expression is conducted for 20 h at 20 °C<sup>13</sup> before harvesting  
222 the cells by centrifugation at 5000  $\times g$  for 30 min at 4 °C. After  
223 washing the cells twice with distilled water and further  
224 centrifugation, pellets were stored at –80 °C. In the case of  
225 *hsTK*, preculture and expression were conducted in dYT  
226 medium (16 g L<sup>-1</sup> tryptone, 10 g L<sup>-1</sup> yeast extract, 5 g L<sup>-1</sup>  
227 NaCl) with 100  $\mu\text{g mL}^{-1}$  ampicillin. Expression was done at 37  
228 °C without induction for 16 h.<sup>21</sup>

229 Bacterial pellets were suspended in 50 mM sodium  
230 phosphate buffer pH 8.0 and NaCl 300 mM (7 mL per  
231 gram) before disruption by sonication. The lysate was  
232 centrifuged at 3220  $\times g$  at 4 °C for 30 min, and the  
233 supernatant was filtered through Whatman #1 paper for  
234 clarification. TKs were purified using a His-trap FF column  
235 connected to an AKTA Start system in 50 mM sodium  
236 phosphate buffer, pH 8.0, 300 mM NaCl. TKs were eluted by  
237 applying an imidazole gradient (80–200 mM, same buffer).  
238 Eluted TKs were dialyzed overnight against 4 L of 5 mM  
239 sodium phosphate buffer pH 8.0 and 30 mM NaCl at 8 °C  
240 using an 8 kDa dialysis membrane (Spectra/Por 6 dialysis  
241 Membrane, Spectrum Lab). Then, TKs were precipitated in  
242 the same dialysis chamber using by 3 M ammonium sulfate in  
243 water at 8 °C for 5 h.<sup>32,33</sup> TKs were stored at 4 °C until used.

244 For analysis, ammonium sulfate was discarded after 2 min  
245 centrifuge at room temperature, and TK pellets were  
246 suspended in 50 mM HEPES buffer pH 7.0, 100 mM KCl  
247 (same volume as the ammonium sulfate sample). Protein  
248 concentration was determined using the Bradford assay with  
249 BSA as the standard. SDS-PAGE was done with an 8%  
250 polyacrylamide stacking gel and a 12% polyacrylamide running  
251 gel. TK monomers were expected to be observed between 69.8  
252 and 77.6 kDa (Table S1).

253 **2.5. Enzymatic Assay.** TK activity was assayed using  
254 Fe(CN)<sub>6</sub><sup>3-</sup> as oxidant of the DHETPP intermediate in the  
255 presence of the sole F6P as ketose (donor) as described by  
256 Kochetov et al. (Scheme 1B).<sup>34</sup> Reaction were performed in  
257 200  $\mu\text{L}$  (final volume) 50 mM HEPES buffer pH 7.0, 100 mM  
258 KCl containing 30  $\mu\text{g}$  of TK, 0.5 mM F6P, 0.5 mM  
259 Fe(CN)<sub>6</sub><sup>3-</sup>, 0.2 mM TPP and 2 mM MgCl<sub>2</sub> ( $n = 3$ ). Kinetics  
260 were triggered by F6P addition, and the decrease of  
261 Fe(CN)<sub>6</sub><sup>3-</sup> absorbance was followed at 420 nm at 37 °C  
262 with a Tristar 5 plate reader (Berthold). A response coefficient  
263 of 472  $\mu\text{M}^{-1}$  per absorbance unit, based on a 0–1.5 mM  
264 Fe(CN)<sub>6</sub><sup>3-</sup> calibration curve, was used to determine the  
265 reaction rates. Specific activities were expressed as  $\mu\text{mol}$  of F6P  
266 converted to E4P per minute per milligram of TK. The  
267 percentage of remaining TPP in purified TKs was determined  
268 by comparing rates with and without the addition of 0.2 mM  
269 TPP.

270 Apparent kinetic parameters were determined by varying  
271 only one substrate concentration, with the others being in  
272 excess. [F6P] was 0–0.5, [Fe(CN)<sub>6</sub><sup>3-</sup>] was 0–2, and TPP was  
273 0–0.2 mM. For *hsTK*, the [F6P] was 0–2 mM. TKs were  
274 always incubated for 15 min with TPP before kinetics.  $K_M^{\text{app}}$   
275 and  $k_{\text{cat}}$  values were determined by fitting data to Michaelis–  
276 Menten and Hanes–Wolf plots. In all experiments, the  
277 conversion was verified to be below 10% to avoid an

underestimation of the reaction rate. Data plotting and kinetic  
analysis (linear and nonlinear fitting) were done using  
Magicplot software (v3.0.1, [www.magicplot.com](http://www.magicplot.com)).

281 **2.6. Homology Modeling.** TKs without described  
282 structure on the PDB were first modeled with Swiss-  
283 MODEL (version of 02/01/2022, <https://swissmodel.expasy.org/>)  
284 before AlphaFold advent.<sup>35</sup> Templates were selected for  
285 modeling by maximizing sequence coverage, GMQE, and  
286 QSQE scores (Table S2). The GMQE evaluates the proximity  
287 of the model obtained to the selected template, while the  
288 QSQE evaluates the quality of the interchain contacts within  
289 the quaternary structures. In agreement with the authors of  
290 Swiss-Model, a good model should have a GMQE  $\geq 0.95$  and a  
291 QSQE  $\geq 0.85$ .<sup>35,36</sup> A QSQE  $\geq 0.70$  can, however, be accepted.

292 Hydrogen atoms were added on each model and on TK  
293 structures available from the PDB using MolProbity v4.5.1  
294 (<http://molprobity.biochem.duke.edu>) in order to include the  
295 volume of hydrogen atoms (~50% of the atoms).<sup>37</sup> Asn, Gln,  
296 and His flips were allowed, and the length of heteroatom-H  
297 bonds was calculated taking account of their electron cloud.  
298 Asn and Gln rotamers were checked with PyMol v2.5.2  
299 (<https://pymol.org>, Schrödinger, LLC) as well as His  
300 tautomers. TK structures available on PDB were also included  
301 in this step and used as a point of comparison to assess the  
302 contribution of hydrogenation and energy minimization in the  
303 models' quality.<sup>38</sup> Energy minimization was performed on  
304 Swiss PDB Viewer v4.10 using the GROMOS96 force field  
305 with 200 steepest descent cycles, 200 conjugate gradient cycles,  
306 and 200 steepest descent cycles for all TK models.<sup>37</sup>

307 The final evaluation of models and structures was carried out  
308 with Verify 3D, which evaluates the overall quality of the  
309 model,<sup>39</sup> ERRAT, which evaluates the local quality of the  
310 model for each residue considering its environment,<sup>40</sup> and  
311 PROCHECK, which evaluates errors on peptide bond angles,  
312 based on the Ramachandran diagram, and of side chains angles  
313 (02.07.2022 version, <https://saves.mbi.ucla.edu/>).<sup>41</sup> All  
314 RMSD were obtained using PyMol by aligning TK structures  
315 or models by pairs with 0 cycles and no transformation (PyMol  
316 RMSD command: align obj1, obj2, cycles = 0, transform = 0).  
317 An additional model was produced using AlphaFold (<https://alphafold.ebi.ac.uk/>)  
318 for *pfTK*.<sup>42</sup> Model evaluation was done  
319 the same.

320 A TK consensus model of the selected TKs based on the  
321 consensus sequence was obtained after multiple sequences  
322 alignment with Clustal  $\omega$  (<https://www.ebi.ac.uk/Tools/msa/clustalo/>)  
323 and determined by JalView v2.11.2.6 (Figures S2  
324 and S3).<sup>43</sup> This sequence was finally modeled with Swiss-  
325 Model using *mtTK* (PDB id 3RIM) as a template.

326 **2.7. Crystallography.** Ammonium sulfate was discarded as  
327 described above. TKs were further desalted by gel-filtration  
328 chromatography in 50 mM HEPES buffer pH 7.0, 100 mM  
329 KCl and concentrated using a 30 kDa cutoff centrifugal device  
330 (Microsep advance, Pall, Saint-Germain-en-Laye, France) at  
331 3220  $\times g$  for 15 min at 4 °C. Then, 1 mM TPP and 2 mM  
332 MgCl<sub>2</sub> (final concentrations) in 50 mM HEPES buffer pH 7.0,  
333 100 mM KCl were added. Final protein concentrations were  
334 measured at 280 nm with a NanoDrop One using the  
335 molecular absorption coefficient predicted by ProtParam  
336 (02.01.2022 version, <https://web.expasy.org/protparam/>)  
337 (Table S1).<sup>44</sup>

338 Crystallization screening conditions were carried out using  
339 the sitting-drop vapor diffusion method and commercial  
340 crystallization kits. For screening, a Mosquito crystallization

341 robot (SPT Labtech) was employed using one protein/  
 342 crystallization agent ratio (100 + 100 nL drops equilibrated  
 343 against 70  $\mu$ L in MRC Crystallization Plates (Molecular  
 344 Dimensions)). Concentrations leading to crystals are provided  
 345 in Table S1 with a growing media different for *spTK* (0.5 M  
 346  $\text{Li}_2\text{Cl}_2$ , 0.1 M Tris pH 8.5, 28% (w/v) PEG 6000), *pfTK* (0.1  
 347 M  $\text{Li}_2\text{SO}_4$ , 0.1 M Sodium citrate pH 5.6, 12% (w/v) PEG  
 348 6000), *hiTK* (0.1 M Magnesium acetate, 0.1 M Sodium citrate  
 349 pH 5.6, 8% (w/v), PEG 10000), *efTK* (0.2 M  $\text{MgCl}_2$  0.1 M  
 350 Tris pH 8.5, 16% (w/v) PEG 4000), and *saTK* (0.1 M tris pH  
 351 8.5, 25% (w/v) PEG 8000). Crystals were further cryopro-  
 352 tected in the same solution, to which 15% ethylene glycol was  
 353 added and harvested. X-ray diffraction data were collected at  
 354 the ID23\_1 beamline (ESRF, Grenoble) at a wavelength of  
 355 0.885600 Å. Data were indexed, integrated, and scaled using  
 356 XDS.<sup>45</sup> The phase problem was solved by molecular  
 357 replacement using phenix.phaser<sup>46</sup> with the herein-constructed  
 358 model. The experimental model was then built using Coot,<sup>47</sup>  
 359 and refinement was carried out using PHENIX.<sup>48</sup> Data  
 360 collection and refinement statistics are listed in Table 2.  
 361 PyMol was used to render the molecular structures.

362 **2.8. Identification of Residues Involved in the Active**  
 363 **Site and Monomer–Monomer Interactions.** Structures of  
 364 TKs crystallized in this work (*efTK*, *hiTK*, *spTK*, *pfTK*,  
 365 *saTK*), those available in the PDB after minimization (*ecTK*,  
 366 *hsTK*, *lmTK*, *mfTK*, *scTK*, *vvTK*<sup>49</sup>), and those modeled here  
 367 (*abTK*, *cgTK*, *mlTK*, *paTK*) were analyzed to identify residues  
 368 involved in the active site and monomer–monomer  
 369 interactions.

370 Residues involved in the active site were obtained using  
 371 CASTp 3.0 (<http://sts.bioe.uic.edu/castp/calculation.html>,  
 372 08.09.2022 version)<sup>50</sup> set with a probe radius of 1.4 Å. The  
 373 cavity corresponding to the active site was selected to identify.  
 374 All TK structures were aligned using PyMol and  $C\alpha$  of a TK  
 375 spatially aligned with  $C\alpha$  of other TK were considered as  
 376 sharing the same spatial position. An additional manual  
 377 completion step was done for  $C\alpha$  not identified by CASTp  
 378 but aligned with the yet identified  $C\alpha$ . This resulted in a  
 379 sequence of spatially aligned residues involved in the active site  
 380 used to construct a representative logo for the TK active site.  
 381 Important residues having the highest number of inter-  
 382 actions are critical for protein activity and folding.<sup>51,52</sup>  
 383 Residues interactomes were obtained from RING 3.0  
 384 (<https://ring.biocomputingup.it/submit>, 08.02.2022 version)  
 385 with the following parameters: “Lollipop” nodes, “multiple”  
 386 edges parameters and “strict” distance threshold (hydrogen  
 387 bond ( $\leq 3.5$  Å), salt bridge ( $\leq 4.0$  Å),  $\pi$ -cation stacking ( $\leq 5.0$   
 388 Å), Van der Waals bond ( $\leq 0.5$  Å),  $\pi$ - $\pi$  stacking ( $\leq 6.5$  Å), and/  
 389 or disulfide bridge ( $\leq 2.5$  Å)).<sup>51</sup> Resulting graphs were  
 390 analyzed with Cytoscape (v3.91.1, <https://cytoscape.org/>).  
 391 Residues having at least 11–12 interactions with neighboring  
 392 residues were selected (14–22% of the residues) to draw up a  
 393 list of important residues (Table S8) and residues interacting  
 394 with residue(s) of the other monomer (as defined by RING)  
 395 to draw up the monomer–monomer interactions list.

396 The three lists of residues (active site, monomer–monomer  
 397 interface, and important residues) were compared to obtain a  
 398 shortlist of residues belonging to these three categories,  
 399 denoted as very important residues (VIRs).

400 The occurrence of residues identified to be part of the active  
 401 site or to the monomer–monomer interface was determined  
 402 (Table S8) and grouped by domains (PP, Pyr, and C-ter). For

Table 2. Data Collection and Refinement Statistics

transketolase	<i>efTK</i>	<i>hiTK</i>	<i>pfTK</i>	<i>saTK</i>	<i>spTK</i>
PDB id data collection	8R3P	8R3O	83RQ	8R3S	8R3R
beamline	ID30B	ID30A-3	PX2A	ID23-1	ID30A-3
wavelength (Å)	0.9762	0.9676	0.9801	0.9781	0.9676
space group	C121	P12 <sub>1</sub>	P12 <sub>1</sub>	P12 <sub>1</sub>	P12 <sub>1</sub>
cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	249.66	110.25	97.04	94.90	93.76
	68.22	90.67	114.34	149.90	119.13
	165.46	155.25	149.90	253.30	136.68
$\alpha$ , $\beta$ , $\gamma$ (deg)	90.00	90.00	90.00	90.00	90.00
	110.71	100.95	90.11	90.20	106.46
	90.00	90.00	90.00	90.00	90.00
resolution range (Å)	34.94–2.90	34.82–2.10	46.81–1.88	34.86–2.80	34.05–2.35
total reflections	158876	565098	1838584	488618	836138
unique reflections	57466	168690	265236	167483	118040
$R_{\text{meas}}$ (%)	17.2 (96.2)	8.4 (20.0)	9.5 (64.6)	26.3 (103.9)	17.6 (97.6)
$CC_{1/2}$ (%)	98.6 (60.7)	99.6 (97.8)	99.8 (84.7)	93.8 (52.1)	99.1 (85.3)
$I/\sigma(I)$	6.41 (1.24)	12.68 (7.29)	17.84 (3.69)	6.74 (1.94)	10.78 (4.07)
multiplicity	2.76 (2.82)	3.34 (3.30)	6.93 (6.02)	2.91 (2.94)	7.07 (7.04)
completeness (%)	98.4	96.3	99.95	96.14	98.50
No. mol./asymm. unit	4	4	4	10	4
refinement					
$R_{\text{work}}/R_{\text{free}}$ (%)	20.94/24.33	15.12/19.03	14.54/16.79	21.58/26.07	15.74/21.57
No. atoms					
protein	20502	20519	21575	50987	20051
ligand/ion	108	155	160	270	116
water	15	2311	2924	0	2363
average B-factor (Å <sup>2</sup> )					
protein	64.77	25.89	18.72	33.25	15.55
ligand/ion	69.14	32.77	27.28	56.83	18.92
water	47.60	33.61	32.79		22.88
R.M.SD.					
bond lengths (Å)	0.005	0.007	0.008	0.004	0.010
angles (deg)	0.804	0.802	0.915	0.672	0.977
Ramachandran					
favoured (%)	97.49	97.85	98.02	96.49	97.06
allowed (%)	2.40	2.15	1.98	3.41	2.94
outliers (%)	0.11	0.00	0.00	0.11	0.00

each position of the consensus sequence, a score was calculated  
 as follows:

$$\text{Score} = \frac{\text{occurrences}}{\text{number of different residues}}$$

where *occurrences* (ranging from 1 to 14) is the number of  
 times any residue was found at a position and *number of*  
*different residues* (ranging from 1 to 20) was the number of  
 different residues at this position. The chemical nature of the  
 side chains of residues was not considered. Residues depicted  
 as letters were placed on a logo regarding their position on the  
 consensus sequence, their height corresponding to their  
 occurrence, and color to their score.

413 **2.9. Dendrogram of TK Evolution.** A more complete  
 414 TKs dendrogram was obtained including 15 new TKs from  
 415 nonpathogenic organisms (*scTK* included) originating from  
 416 plants, bacteria, yeasts, parasites, and animals (Table S9).  
 417 These 30 sequences were aligned using Clustal  $\omega$  with the  
 418 following settings: Output format: ClustalW, Dealign input:  
 419 No, MBED-Like clustering: Yes, MBED-Like clustering during  
 420 iteration: Yes, Number of combined iterations: default (0),  
 421 Maximum guide tree iteration: default, Maximum HMM  
 422 iterations: default, Output distance matrix: yes, Output guide  
 423 tree: yes, Order: input. The alignment was then annotated by  
 424 using JalView. The dendrogram was constructed using iTOL  
 425 (<https://itol.embl.de/>), with divergences from Clustal  $\omega$  as  
 426 distances (calculated based on the alignment score). A  
 427 normalized distance of 1 was assigned to the two most  
 428 divergent sequences (*Plasmodium falciparum* and *Schistosoma*  
 429 *mansoni* TKs), positioning the dendrogram origin equidistant  
 430 between these two sequences. Other TK sequences were  
 431 placed accordingly.

### 3. RESULTS

432 **3.1. TKs Expression and Purification.** Fourteen TKs are  
 433 expressed in *E. coli* BL21(DE3) and purified by Ni-NTA  
 434 affinity chromatography. Purification yields, quantities, and  
 435 specific activities differ regarding the enzyme (Table S1). Not  
 436 surprisingly, bacterial TKs such as *abTK*, *efTK*, *ecTK*, and  
 437 *vvTK* are well expressed in *E. coli* (>90 mg per 2 L culture,  
 438 purification yields >25%). Higher eukaryotic TKs such as  
 439 *lmTK* and *hsTK* are far less expressed and purified (<20 mg  
 440 and yields <3% for a 2 L culture). The latter TKs require  
 441 different expression protocols: *hsTK* is expressed in dYT  
 442 medium without induction at 37 °C while *lmTK* is expressed  
 443 in LB medium using 0.4 mM IPTG for induction at 15 °C.<sup>21,53</sup>  
 444 Surprisingly, *mtTK* and *mlTK* that share high sequence  
 445 identity (>89%) are expressed with different yields. All TKs  
 446 are recovered and have the expected molecular weight as  
 447 shown by SDS-PAGE with few contaminants, with the  
 448 exception of *paTK* that is copurified with another protein of  
 449 about 100 kDa (Figure S1; Table S1).

450 After dialysis and precipitation in ammonium sulfate, less  
 451 than 11% of the active site are found to be occupied by TPP  
 452 allowing to determine the apparent Michaelis constant for TPP  
 453 ( $K_{\text{TPP}}^{\text{app}}$ ) except for *hsTK*, *mtTK*, and *pfTK* with more than  
 454 30% of the active sites occupied by TPP (Table S1). All TKs  
 455 are active using F6P as the donor and  $\text{Fe}(\text{CN})_6^{3-}$  as the  
 456 oxidant of the DHETPP intermediate (Scheme 1B). Specific  
 457 activities are slightly higher for prokaryotic TKs ( $0.117 \pm$   
 458  $0.043 \mu\text{mol min}^{-1} \text{mg}^{-1}$ ) compared to eukaryotic TKs ( $0.073$   
 459  $\pm 0.056 \mu\text{mol min}^{-1} \text{mg}^{-1}$ ) using this assay.

460 **3.2. Kinetic Parameters.** Kinetic parameters (e.g.,  
 461 apparent Michaelis constants for TPP, F6P,  $\text{Fe}(\text{CN})_6^{3-}$ , and  
 462  $k_{\text{cat}}$  values) are determined for 11 TKs (Table 3).  $K_{\text{TPP}}^{\text{app}}$   
 463 values are in the micromolar range, as previously observed for  
 464 other TKs. This very high affinity between TKs and its cofactor  
 465 explains why it is often difficult to remove the TPP from the  
 466 active site as observed for *hsTK*, *mtTK*, and *pfTK*.<sup>21,54</sup>  
 467 Numerous studies have focused on the use of TPP-derived  
 468 molecules to inhibit TKs (2'-methoxy-thiamine, oxythiamine,  
 469 4-anilinoquinoline triazine) in particular in oncology.<sup>5,55,56</sup>  
 470 Taking into account the  $K_{\text{TPP}}^{\text{app}}$  of these TKs, it would be  
 471 more pertinent to look for inhibitors that bind elsewhere on  
 472 TKs rather than TPP competitors.<sup>5,49</sup>

**Table 3. Apparent Kinetic Parameters ( $K_{\text{F6P}}^{\text{app}}$ ,  $K_{\text{FeCN6}}^{\text{app}}$ ,  $K_{\text{TPP}}^{\text{app}}$ , and  $k_{\text{cat}}$ ) for 11 TKs<sup>a</sup>**

TK	$K_{\text{F6P}}^{\text{app}}(\mu\text{M})$	$K_{\text{Fe}(\text{CN})_6}^{\text{app}}(\mu\text{M})$	$K_{\text{TPP}}^{\text{app}}(\mu\text{M})$	$k_{\text{cat}}(\text{s}^{-1})$
<i>abTK</i>	182.1 ± 4.2	693.8 ± 7.1	0.48 ± 0.01	1.68 ± 0.05
<i>cgTK</i>	172.0 ± 3.4	702.2 ± 10.7	1.13 ± 0.01	0.37 ± 0.01
<i>efTK</i>	169.1 ± 5.0	331.5 ± 5.2	4.39 ± 0.09	0.38 ± 0.02
<i>ecTK</i>	298.5 ± 4.9	1009.2 ± 18.4	1.56 ± 0.02	0.22 ± 0.01
<i>hrTK</i>	123.1 ± 1.0	586.0 ± 16.3	1.68 ± 0.01	0.22 ± 0.01
<i>hsTK</i>	443.9 ± 17.6	1385.2 ± 9.2	n.d.	0.06 ± 0.01
<i>mtTK</i>	237.4 ± 8.7	595.3 ± 13.7	n.d.	0.58 ± 0.04
<i>pfTK</i>	128.2 ± 1.4	1069.7 ± 5.1	n.d.	0.72 ± 0.03
<i>saTK</i>	253.8 ± 9.3	789.3 ± 16.6	2.36 ± 0.04	0.19 ± 0.01
<i>spTK</i>	78.0 ± 2.8	468.0 ± 4.1	1.86 ± 0.01	0.14 ± 0.01
<i>vvTK</i>	154.7 ± 1.8	254.3 ± 5.8	1.99 ± 0.25	0.20 ± 0.01
Average	203 ± 96	716 ± 318	1.97 ± 1.14	0.43 ± 0.43

<sup>a</sup>Rates were obtained in 50 mM HEPES buffer pH 7.0, 100 mM KCl, 2 mM  $\text{MgCl}_2$ , pH 7.0 at 37 °C containing 0–500  $\mu\text{M}$  F6P, 0–200  $\mu\text{M}$   $\text{Fe}(\text{CN})_6^{3-}$ , and 0–200  $\mu\text{M}$  TPP ( $n = 3$ ). Only one substrate concentration is varied for the determination of the Michaelis–Menten constants, with the others being at maximal concentration.

The  $K_{\text{F6P}}^{\text{app}}$  values are in the 100–200  $\mu\text{M}$  range with a  
 slight difference between bacterial TKs ( $187 \pm 72 \mu\text{M}$ ) and  
 lower eukaryotes TKs (*pfTK* and *cgTK*, 128 and 172  $\mu\text{M}$ ,  
 respectively). These values agree with the  $K_{\text{F6P}}^{\text{app}}$  values  
 determined by Kochetov et al. for *hsTK* (340  $\mu\text{M}$ ), but Fullam  
 et al. found a higher value for *mtTK* (600  $\mu\text{M}$ ) using the same  
 assay but with a lower  $\text{Fe}(\text{CN})_6^{3-}$  concentration (0.5  
 mM).<sup>13,22</sup> Slightly higher  $K_{\text{F6P}}^{\text{app}}$  value is obtained for *hsTK*,  
 the only higher eukaryote TK of this study, and could reflect  
 the specialization of the enzyme during evolution. From a  
 metabolic point of view, reactions of the nonoxidative part of  
 the PPP aim to produce F6P, despite these reactions being  
 equilibrated. Then, it is not totally astonishing that TKs from  
 higher species are not prone to use F6P as substrate (from a  
 kinetic point of view) to orient metabolic flux in the desired  
 direction, on the contrary to bacterial TKs that are more  
 tolerant to the substrate range.

The second substrate,  $\text{Fe}(\text{CN})_6^{3-}$ , is not a natural substrate  
 for TKs, but our recent work shows that it behaves as a  
 Michaelian one.<sup>49</sup> The rates versus  $[\text{Fe}(\text{CN})_6^{3-}]$  plots are  
 hyperbola (Figure S7; Table S9) confirmed by linear  
 representations of the data (Hanes–Wolf plots, data not  
 shown). Therefore, it is possible to determine the  $K_{\text{Fe}(\text{CN})_6}^{\text{app}}$   
 values for all TKs. The  $K_{\text{Fe}(\text{CN})_6}^{\text{app}}$  values are from 254  $\mu\text{M}$  for  
*vvTK* to more than 1380  $\mu\text{M}$  for *hsTK*. The average  $K_{\text{Fe}(\text{CN})_6}^{\text{app}}$   
 value for bacterial TKs is  $590 \pm 245 \mu\text{M}$  while it is  $1052 \pm 342$   
 $\mu\text{M}$  for eukaryotic TKs suggesting that more evolved TKs  
 (*pfTK*, *cgTK*, and *hsTK*) and do not accept easily non-natural  
 substrates, such as the bulky  $\text{Fe}(\text{CN})_6^{3-}$  anion.

The catalytic constants ( $k_{\text{cat}}$ ) are in the range of 0.14–1.68  
 $\text{s}^{-1}$  for all TKs except for *hsTK* which has a very low  $k_{\text{cat}}$  value  
 ( $0.06 \pm 0.01 \text{ s}^{-1}$ ) with this assay, as reported in the literature.<sup>14</sup>  
 Overall, these TKs do not differentiate strongly regarding their  
 kinetic parameters with the exception of *hsTK*.

**3.3. Molecular Modeling of TKs.** Several publications  
 describe the role of specific conserved histidyl residues in  
*ecTK*, *hsTK*, and *scTK* involved in TPP stabilization,<sup>24,57,58</sup>  
 activation,<sup>14,15,19,20</sup> and catalytic substrate stabiliza-  
 tion.<sup>3,16,24,57,59</sup> To identify which conserved residues are  
 essential for activity and understand which residues may  
 explain the catalytic differences observed, we carried out a

514 comparative study of (i) the residues present in the active site  
515 of TKs in our panel and (ii) residues involved in the  
516 monomer–monomer interactions. As structures of expressed  
517 TK are not all available (or incomplete like *pa*TK, PDB id  
518 4XEU), molecular models are constructed based on available  
519 structures (Table 1). The structure of the active site of *pa*TK is  
520 incomplete (PDB id 4XEU). All enzymes are modeled by  
521 homology using Swiss Modeler and AlphaFold once it  
522 becomes easily available, as described in the experimental  
523 part (Table S2).

524 TK models have more than 60% sequence identity with their  
525 template, except for *pf*TK, which has only 46% sequence  
526 identity with the TK from *Scheffersomyces stipitis*. Four TK  
527 models are highly confident ( $GMQE \geq 0.95$ ) and five other  
528 models are acceptable ( $0.90 \leq GMQE < 0.95$ ). The only  
529 model showing insufficient quality ( $GMQE = 0.82$ ) is the  
530 *pf*TK model. Interchain contacts are properly modeled for all  
531 TKs ( $QSQE \geq 0.85$ ) except for *cg*TK closed to the threshold  
532 value ( $QSQE = 0.84$ ). The average  $\Delta G^\circ$  of the models  
533 decreases from  $-50.9 \pm 3.9 \times 10^3$  kJ mol<sup>-1</sup> before  
534 hydrogenation and minimization to  $-76.0 \pm 3.3 \times 10^3$  kJ  
535 mol<sup>-1</sup> after these steps (Table S2). The  $\Delta G^\circ$  of TK structures  
536 from the PDB decrease from  $-48.7 \pm 4.9 \times 10^3$  to  $-75.1 \pm 1.2$   
537  $\times 10^3$  kJ mol<sup>-1</sup> after minimization. All TK models have a 3D/  
538 1D score higher than 95% using Verify 3D confirming the  
539 good quality of the models. Again, the *pf*TK model is below  
540 this indicative value (3D/1D score = 92.26%) as well as the  
541 *hs*TK structure available from the PDB (3D/1D score =  
542 93.02%). High ERRAT scores ( $\geq 95\%$ ) are obtained for *hi*TK  
543 while TK structures present on the PDB and six other models  
544 have a score between 90% and 95%. It is worth noting that  
545 experimental structures obtained by X-ray crystallography have  
546 ERRAT scores below 95%. Therefore, scores higher than 90%  
547 correspond to acceptable models and structures. The *sp*TK  
548 model has an almost acceptable ERRAT score (89.98%), but  
549 the *pf*TK is far below this threshold (87.12%).

550 Ramachandran diagrams indicate that all TKs have residues  
551 with peptide bond angles in acceptable regions (more than  
552 99%). Among these residues, 85.2–89.5% are predicted to be  
553 in favorable regions (86.6% for *hs*TK available on the PDB).  
554 The majority of poorly modeled residues are in regions in  
555 contact with the solvent, especially in intrinsically unstructured  
556 regions or loops (data not shown).

557 Although *pf*TK model does not reach the quality required  
558 for use in the resolution of crystallographic structure, it is of  
559 higher quality than the one proposed by Anayet Hasan et al.  
560 using Modeler v9.13 software and *Zea mays* TK (PDB id  
561 1ITZ) as template (3D/1D score = 0.77, ERRAT score =  
562 78.31, residues predicted in favored and allowed regions =  
563 99.5%).<sup>60</sup> This underlines the difficulty in modeling *pf*TK,  
564 especially over large portions of the C-ter domain (Table S3).

565 The modeling protocol is validated by comparison of TK  
566 structures available on the PDB with the corresponding model  
567 after hydrogenation and energy minimization (Table S7). The  
568 average RMSD is  $0.127 \pm 0.035$  Å, indicating a very low  
569 impact of hydrogenation and energy minimization on the  
570 structure of the TKs obtained from the PDB. Hydrogenation  
571 of the structure before energy minimization allows to refine the  
572 position of the side chains without modifying the main chain.  
573 The threshold values for model validation are all verified  
574 (Verify 3D score  $\geq 95.0\%$ , ERRAT  $\geq 90.00\%$ , residues in  
575 allowed of favored region of the Ramachandran diagram  
576  $\geq 99.0\%$ ) or closed to them (*sa*TK, *sp*TK).

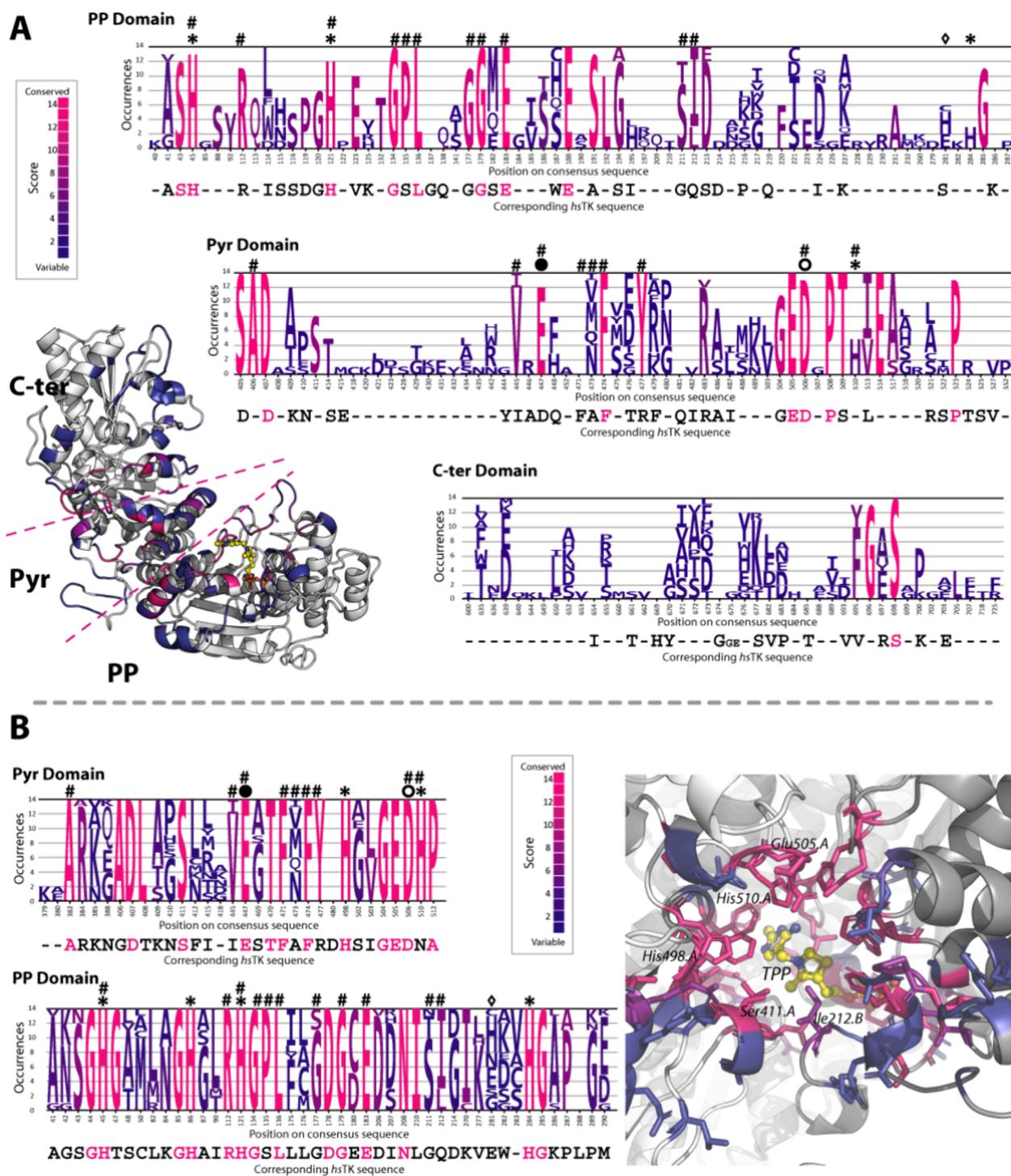
**3.4. TK Crystallography.** Five new structures of TK are 577  
obtained by X-ray crystallography (*ef*TK, *hi*TK, *pf*TK, *sa*TK, 578  
and *sp*TK) (Table 2). The electron density maps of these TKs 579  
are all solved by molecular replacement using corresponding 580  
models, except for *pf*TK the structure of which is solved using 581  
an AlphaFold model available on the UniProt (AF-C6KSV3- 582  
F1). Divergences still exist in the C-ter domain in both models 583  
(Tables S3 and S4).<sup>42</sup> 584

These TKs are homodimers in interlocking V-shaped chains 585  
as all other described TK.<sup>14,15</sup> Each monomer in these new 586  
structures is formed by PP, Pyr, and C-ter domains connected 587  
by linker I and linker II as classically observed for all described 588  
TKs (Tables S5 and S6). All TKs have two identical active sites 589  
located at the monomer–monomer interface, formed by the 590  
PP domain of one chain and the Pyr domain of the other 591  
chain. In general, TK structures appear to be particularly well 592  
conserved, as shown by the secondary structures forming each 593  
domain. PP domains are formed by 10–12  $\alpha$  helices, 5  $\beta$  594  
strands, and 14–17 loops, the Pyr domain is formed by 7–9  $\alpha$  595  
helices, 5–6  $\beta$  strands, and 9–12 loops, and the C-ter domain 596  
is formed by 4–5  $\alpha$  helices, 5  $\beta$  strands, and 7–11 loops. While 597  
linker II is nearly always formed by 1  $\alpha$  helix and 2 loops 598  
(intrinsically unfolded), linker I is formed by 1–3  $\alpha$  helices and 599  
1–4 loops. Some particularities are nevertheless observed: 600  
*sp*TK is unique because it has a  $\beta$ -strand in each of these two 601  
linkers, whereas *mt*TK has only one in linker II. These 602  
observations confirm that *hs*TK is still unique among known 603  
TK as its PP domain is composed of only 8  $\alpha$ -helices and is the 604  
only TK to have fully unstructured linkers. These structures 605  
and models allow the building of a consensus sequence (Figure 606  
S2) and a consensus structural model (Figure 1A). 607

**3.5. Ubiquitous Histidine Crown.** A ubiquitous motif is 608  
identified: a crown of 6 histidyl residues (5 in *hs*TK) 609  
surrounding the TPP binding site closed to the ketose binding 610  
site (Figure 1B, Table S9) yet observed for the *vv*TK.<sup>49</sup> A 611  
seventh less systemic histidyl residue is also identified (His 281 612  
in the consensus sequence). The literature only describes the 613  
role of some of these residues in *ec*TK and *sc*TK such as 614  
stabilizing the interactions between TK and TPP via 615  
pyrophosphate group (His 86, His 284, and His 498 on the 616  
consensus TK) and thiazolium ring (His 617  
510).<sup>3,14,15,24,54,55,58,59,61,62</sup> This crown is also involved in 618  
TPP activation by favoring the deprotonation of the thiazolium 619  
ring by the N4' of the aminopyrimidine ring of TPP (His 45, 620  
His 121, and His 510) and in activation by Glu 447 (Figure 621  
S5).<sup>15,19,20</sup> His 45, His 121, and His 510 are close to the 622  
conserved Asp 506, forming a subnetwork of four residues and 623  
stabilizing the proton released from TPP activation.<sup>63,64</sup> 624

*hs*TK differs from other TKs because it has only five histidyl 625  
residues in the crown: His 510 in consensus TK is substituted 626  
by a Gln (residue 428 in the *hs*TK sequence).<sup>14</sup> This residue 627  
seems to have a minor role in interaction with TPP as  $K_{TPP}^{app}$  628  
is similar for all TKs (Table 3). However, within *hs*TK, the 629  
proton resulting from TPP activation could have a more 630  
transient position between the imidazole rings of His 45 and 631  
His 121 and the C2 carbanion of the thiazolium ring of TPP, 632  
which could explain why *hs*TK has the lowest activity. 633

**3.6. Monomer–Monomer Interface and Active Site** 634  
**Mapping.** Alignment of TK models and X-ray structures 635  
allows mapping of the residues forming the monomer– 636  
monomer interface and residues forming the active site. 637  
Unique residues (e.g., residues that are found in only one TK 638  
at a unique spatial position) are used to characterize the 639



**Figure 2.** Monomer interface mapping (A) and homodimer active site mapping (B) of the 14 TKs (Table 1 plus *scTK*, Table S7). Logos show residue occurrences for each TK consensus position, with scores determined as described in the experimental section, from dark blue (variable residues) to pink (conserved residues) (Figure S2). The *hsTK* sequence is aligned below the logos, and residues identical to those having a score  $\geq 12$  are colored in pink. Residues of the histidine crown (\*), His 281 ( $\diamond$ ), Glu 447 initiating TPP activation ( $\bullet$ ), Asp 506 stabilizing the thiazolium ring of the protonic TPP (O) and VIRs (#) are indicated. If no residues are shown on the logos, only *hsTK* has a residue at this spatial position. Residues are projected onto the monomeric (A) or dimeric (B) structure of the consensus TK model (white or white and gray ribbons, respectively) using the same color as for the logos. TPP is represented by balls and sticks (carbon is yellow, nitrogen blue, oxygen red, and phosphorus orange). The three domains are separated by pink dotted lines (A). Some conserved residues are also shown as sticks and labeled for ease of orientation (B).

640 specificities of TKs (Table S7). Monomer–monomer inter-  
 641 faces in TKs involve 79–95 residues from all domains except  
 642 linker I (Figure 2A, Table S8). The nature of the residues at a

spatial position is highly variable. The interfaces of prokaryotic 643  
 TKs are not specific with only 2–16 unique residues (2.4–17% 644  
 of the interface interactome), while TKs of lower eukaryotes 645

646 are slightly more specific with 12–22 unique residues (13–  
647 26%, Table S7). In contrast, *hsTK* has a highly specific  
648 interface formed by 54 unique residues over 89 (60%).

649 Main differences in residues forming the interface are found  
650 from residues 214 to 280 (two  $\alpha$ -helices and two  $\beta$ -strands of  
651 the PP domain), 418 to 442 (loop on the Pyr domain), and  
652 600 to 662 (two helices of the C-ter domain) of the consensus  
653 sequence where residues of TKs from pathogens have no  
654 equivalent in *hsTK* (Deep blue residues in Figure 2A).  
655 Conserved residues in this interface are mostly found in the N-  
656 terminal region of the PP domain (43–192), which  
657 corresponds to residues also involved in the active site, and  
658 in the C-terminal part of the Pyr domain (404–428). Residues  
659 Glu 447, Asp 506, and the histidine crown take part in the  
660 monomer–monomer interaction with the exception of His 86  
661 and His 510. Interestingly, Glu 505, a neighbor of Asp 506, is  
662 highly conserved and could also be involved in proton  
663 stabilization resulting from TPP activation (Pyr domain log;  
664 Figure 2A).

665 The number of residues forming the active sites is a nearly  
666 constant number over all TKs (78–82 residues, Figure 2B,  
667 Table S7) and appears to be conserved regarding the nature of  
668 residues at a given position. All histidyl residues belonging to  
669 the crown as well as Glu 447, Glu 505, and Asp 506 are  
670 conserved within all TKs, confirming the importance of these  
671 residues. The number of unique residues in the active site is  
672 fairly variable (0–13 residues), except for the *hsTK* active site  
673 with 33 unique residues. Therefore, *hsTK* has a nearly unique  
674 active site and monomer–monomer interface compared to  
675 those of other TKs.

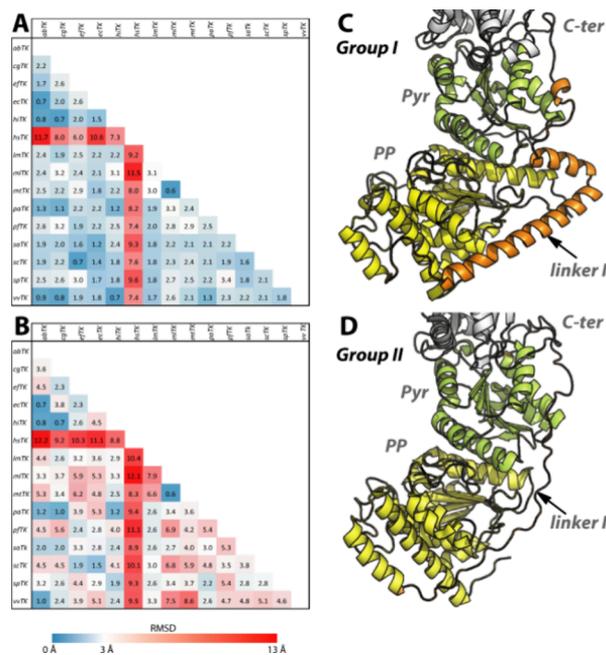
676 Depending on the TK, 80–133 important residues per  
677 monomer are identified (Table S8). They are mostly found in  
678 the core protein, are not accessible to the solvent, and have an  
679 average of 13 interactions per residue (11–19). This  
680 represents less than 16% of the residues of TKs (662–680  
681 residues), except for the longest *mTK* and *mtTK* (17%, 699–  
682 700 residues) and *hiTK* (19.8%, 665 residues). In contrast,  
683 *hsTK*, although the smallest TK, has 133 important residues  
684 (21.2% of 623 residues). Despite the fact that *hsTK* is shorter  
685 than the other TKs, its interactome is highly developed.

686 After comparing the important residues to those belonging  
687 to the active site and the monomer–monomer interface, a  
688 condensed list of crucially accessible residues, referred to as  
689 VIRs, is now accessible (Table S8). The number of VIR  
690 comprises between 3 and 18 (17 for *hsTK*) considering both  
691 chains. Nine of them are unique in *hsTK* while they are  
692 conserved in other TKs (3 unique VIR at the maximum). Once  
693 again, this indicates that *hsTK* has unique and more numerous  
694 residues in the critical positions of the dimer interface and in  
695 its active site.

696 Residues of the histidine crown (His 45, His 121, His 510)  
697 and residues involved in TPP activation (Glu 447, Asp 506)  
698 are not systemic VIRs (Table S8). On the contrary, Glu 505 is  
699 probably not a VIR because it is deeply buried in the active site  
700 and not accessible to the solvent once the TPP is present.

701 **3.7. Structural Conservation of TKs and Differences.**

702 TKs share an evolutionarily conserved structure, but the *hsTK*  
703 differs in some ways. Alignment matrices based on the  
704 comparison of RMSD between monomers and between dimers  
705 are constructed to quantify these structural differences (Figure  
706 3). As a preliminary analysis, RMSD values between each  
707 monomer of a single TK are determined to be close to 1 Å.  
708 This ensures that monomers within a dimer are structurally

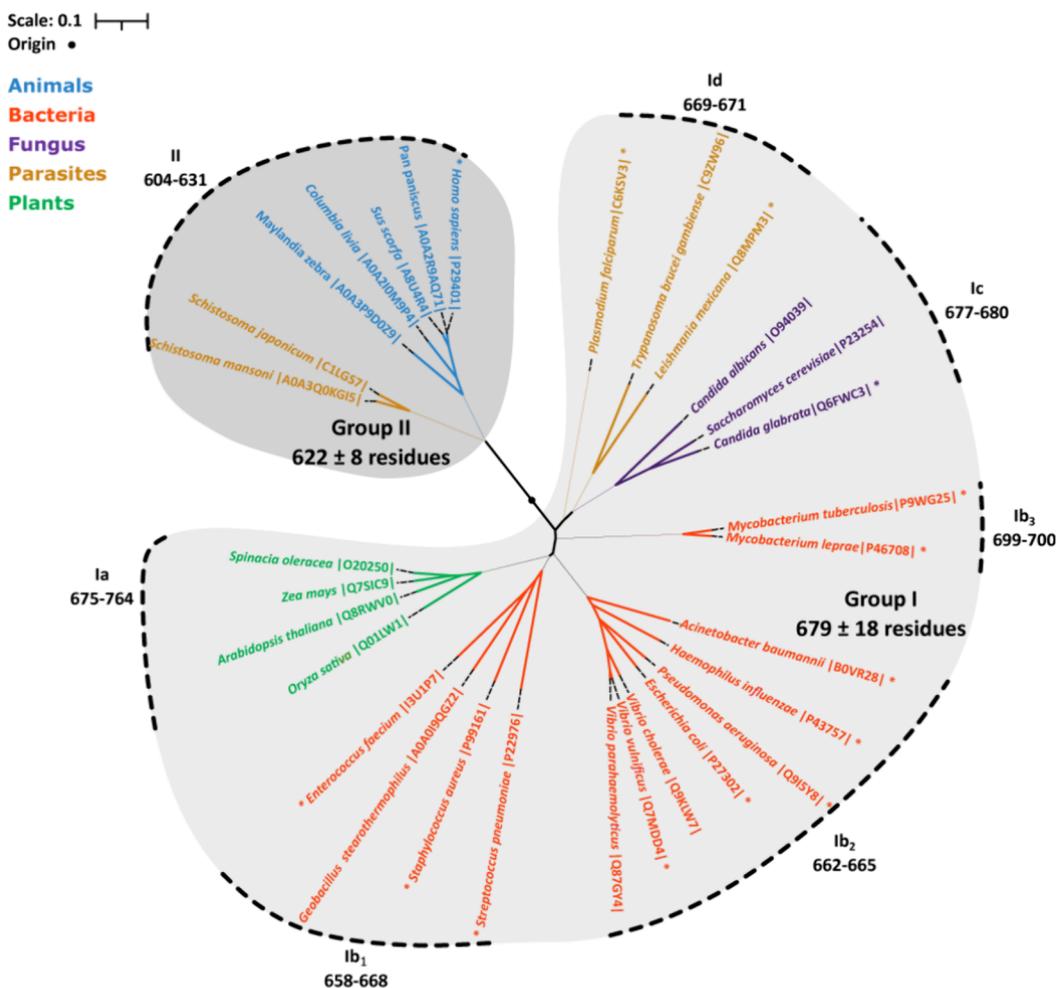


**Figure 3.** RMSD calculated between TKs monomers (A) and TK homodimers (B). RMSD is expressed as a color scale from blue to red (from 0 to 13 Å). Representative structure of TK monomers of group I (C) and of group II (D) highlighting linker I in orange. PP, Pyr, and C-ter domains are in yellow, green, and white, respectively.

equivalent, and further calculations are not biased (Table S12).  
709 Comparison of monomers of all TKs shows that they are  
710 highly similar (RMSD < 3.5 Å) except *hsTK* which exhibits an  
711 RMSD value ranging from 6.0 to 11.750 Å (Figure 3A). This  
712 confirms quantitatively that the tertiary structure of monomers  
713 of *hsTK* is different from the other monomers. This is  
714 reinforced by comparing the RMSD values of TK homodimers,  
715 all sharing at least one RMSD value below 3 Å with another  
716 TK, ensuring continuity between their structures. As expected,  
717 *hsTK* has RMSD values higher than 8.3 Å with any of the other  
718 14 TKs. However, its place in TK evolution is questionable as  
719 the secondary structures of domains are globally conserved.  
720

**3.8. TK Evolution.** In order to understand the position of  
721 *hsTK* within this enzyme evolution, sequence selection is  
722 expanded to other TKs, not the WHO lists (*vide infra*)  
723 including sequences from plants, animals, nonpathogenic  
724 bacteria, parasites, and fungi. The 30 aligned sequences are  
725 represented as a dendrogram (Figures 4, S3, and S6; Table  
726 S11).  
727

TKs are divided into two groups, regarding their average size  
728 and then into subgroups of the clade to which they belong.  
729 The first group of TKs is 658–741 residues long and  
730 corresponds to TKs of plants (subgroup Ia), bacteria  
731 (subgroups Ib<sub>1</sub>, Ib<sub>2</sub>, and Ib<sub>3</sub>), fungi (subgroup Ic), and  
732 parasites (subgroup Id). All of these sequences are well  
733 distributed between the different clades. Parasites (without  
734 *Schistosoma* sp.) and fungi clades are homogeneous with  
735 669–671 and 677–680 residues long TKs, respectively. Plants TKs  
736 are more heterogeneous with sequences ranging from 675 (*Zea*  
737 *mays*) to 741 residues (*Arabidopsis thaliana* and *Spinacia*  
738 *oleracea*). The bacteria clade is homogeneous with sequences  
739 ranging from 658 to 668 (subgroups Ib<sub>1</sub> and Ib<sub>2</sub>) residues with  
740 a clear discrepancy for *Mycobacterium* TKs (699–700 residues,  
741 subgroup Ib<sub>3</sub>) that seems to diverge earlier than other bacteria  
742



**Figure 4.** Dendrogram representing TK evolution. The 30 TK sequences (Table S11) are shown here after multiple alignment of their primary sequences using Clustalw  $\omega$  (Figure S3). Animal TKs are colored in blue, bacterial TKs in red, fungal TKs in purple, parasite TKs in yellow, and plant TKs in green. Origin and scale are set as defined in the experimental part. UniProt accession numbers are given for each TK, and those discussed earlier in this study are marked with a star (\*).

743 as observed by the conservation of additional residues. From  
 744 an evolutionary point of view, they are closer to TKs of fungi  
 745 and parasites than to bacteria, which means that the  
 746 *Mycobacterium* TK clade could diverge from one of the oldest  
 747 common ancestors in this study. The second group of TKs  
 748 diverges significantly from the first group with TKs of shorter  
 749 size (604–631 residues). It is formed by TKs from animals and  
 750 from *Schistosoma* sp. It is possible that horizontal gene transfer  
 751 occurred between their common ancestors because of the  
 752 host–parasite relationship. All TKs in the second group would  
 753 be structurally closer to *hsTK* than to TKs of the first group  
 754 because sequences are shorter, in particular in the region  
 755 corresponding to linker I compared to the first group of TKs  
 756 (Figure S3, sequences highlighted in pink). Furthermore, a  
 757 loop previously considered to be conserved within TK  
 758 evolution finally differs slightly between the first and the  
 759 second group.<sup>15</sup> The residues of the histidine crown involved  
 760 in TPP activation, such as Glu 447 and Asp 506 (Table S8;  
 761 Figure S3), are conserved in both groups. The only exception  
 762 is His 510 (in the consensus sequence) replaced by a Gln in  
 763 Group II.  
 764 AlphaFold models (or X-ray structure, if available) are  
 765 collected for all TK of the dendrogram. Domain analysis  
 766 highlights that TKs of group I have a linker I of  $77 \pm 2$  residues

while TK of group II have all a sorter linker I of  $40 \pm 2$   
 residues (Figure 3; Tables S5 and S6). RMSD calculation  
 between all TK monomers is below 1.9 Å within TK of group  
 II, suggesting they all have the same tertiary structure, while  
 RMSD values between TKs of group I and TKs of group II are  
 between 4.9 and 14.2 Å (Figure S6). This clearly confirms that  
 two TK groups exist, mainly differing by linker I size and  
 secondary structure. Within group I, RMSD values between  
 monomers are below 3.6 Å s (and below 2.1 Å within the  
 subgroup), showing a conservation of the fold in this group.  
 RMSD calculation for TKs of plants should be considered with  
 care as their PP domain is significantly longer than for the  
 other TKs and is not properly predicted in AlphaFold models.

#### 4. DISCUSSION

The aim of this study is to decipher the differences and  
 similarities among all TKs and understand the place of *hsTK*  
 in the evolution of this enzyme. Among the 14 TKs identified  
 in the first set of enzymes, 11 are expressed and purified, and  
 their apparent kinetic constants are obtained. Prokaryotic TKs  
 are more active than eukaryotic TKs (with the exception of *pfTK*)  
 using F6P and  $\text{Fe}(\text{CN})_6^{3-}$  as substrates. This could be linked  
 to the fine regulation of the PPP in eukaryotes whereas  
 prokaryotes require a very active PPP in order to multiply

789 quickly. Furthermore, substrate affinities may differ from one  
790 TK to another for F6P, as evidenced by the  $K_{F6P}^{app}$  values for  
791 TKs of group I (78–298  $\mu$ M, Table 3) and for *hs*TK (444  $\mu$ M,  
792 group II).

793 For the first time, the X-ray structures of TKs from *E.*  
794 *faecium*, *H. influenzae*, *P. falciparum*, *S. aureus*, and *S.*  
795 *pneumoniae* are obtained. To solve the structures and  
796 compensate for the lack of structure for *ab*TK, *cg*TK, *ml*TK,  
797 and *pa*TK, homology modeling is carried out with a protocol  
798 that allows reliable models for TKs belonging to the same clade  
799 (Table S2). The model based on the Swiss Model for *pf*TK is  
800 of better quality than the one described in 2007 by Hasan et  
801 al.<sup>60</sup> but is not sufficient to resolve the electron density map.  
802 For *pf*TK, which clade is not yet represented in the PDB,  
803 modeling using AlphaFold is preferred, and a new structure of  
804 *pf*TK is obtained extending the knowledge of this enzyme  
805 expressed for the first time in 2007.

806 All TKs of this work share ubiquitous structural and  
807 functional features including *hs*TK. The three domains,  
808 organized as Rossmann folds, show few variations regarding  
809 the number of  $\alpha$ -helices,  $\beta$ -strands, and length of individual  
810 domains. We identified a crown of six histidyl residues placed  
811 around TPP except for *hs*TK, which has only five. The role of  
812 some of them in the stabilization and activation of TPP has  
813 been previously identified in *ec*TK and  
814 *sc*TK.<sup>3,14,15,24,54,55,58,59,61,62</sup> Two additional residues involved  
815 in TPP activation (Glu) and proton stabilization (Asp) are also  
816 conserved. Two of the histidyl residues of the crown are  
817 systematically involved in the interaction between the  
818 monomers. The present work proposes extending their  
819 importance for all TKs. From the TK evolutionary point of  
820 view, TKs of this work belong to the group composed of TKs  
821 from plants, bacteria, fungi, and parasites (group I), while  
822 *hs*TK belongs, with animals and *Schistosoma sp.*, to another  
823 group (group II). Both groups differ by the length and folding  
824 of linker I, mostly formed by helices and  $77 \pm 2$  residues long  
825 in group I, while it is nonstructured and  $40 \pm 2$  residues long  
826 in group II (Figure 3). The two first helices of linker I in TK of  
827 group I form a V-shape motif interacting with a loop of the PP  
828 domain (residues 157–167 of the consensus sequence). This  
829 motif closes a pocket proposed to be a cooperative inhibitor  
830 binding site in *vt*TK.<sup>49</sup> In TK of group II, these motifs are  
831 absent, and the pocket is not closed by linker I. Moreover,  
832 *hs*TK is not inhibited by this inhibitor (*unpublished data*). It  
833 could be speculated that regulation would be different for TK  
834 of both groups and that drugs targeting this motif would be  
835 efficient against human pathogens without affecting the host.

836 The *hs*TK has the highest number of unique residues in the  
837 monomer–monomer interaction region and in its active site  
838 compared to TKs from pathogens of group I. Moreover, *hs*TK,  
839 despite its smaller size, maximizes interaction number to  
840 maintain its dimeric structure. Considering the importance of  
841 PPP in cellular metabolism and the difference between TKs of  
842 groups I and II, it is conceivable in the future to identify TK  
843 inhibitors used as antibiotics or in chemotherapies.<sup>5,9,10,55,56,65</sup>  
844 Thanks to new knowledge, notably through mapping, it would  
845 be possible to inhibit more specifically the TKs of human  
846 pathogens by specifically targeting VIRs and/or monomer–  
847 monomer interfaces.

■ ASSOCIATED CONTENT	848
④ Supporting Information	849
The Supporting Information is available free of charge at	850
<a href="https://pubs.acs.org/doi/10.1021/acs.biochem.3c00714">https://pubs.acs.org/doi/10.1021/acs.biochem.3c00714</a> .	851
Consensus sequence, modeling parameters, sequence	852
alignments, purification table, domains size, RMSD	853
values, activation mechanisms, and Michealis–Menten	854
plots (PDF)	855
Reaction rates (XLSX)	856
Monomer–monomer interaction, active site, important	857
and very important residues (XLSX)	858
PDB structures from X-structures, PSE consensus	859
model, and PDB of TK models (ZIP)	860

■ AUTHOR INFORMATION	861
Corresponding Authors	862
Franck Charmantray – Université Clermont Auvergne, CNRS,	863
SIGMA Clermont, Institut de Chimie de Clermont-Ferrand	864
(ICCF), F-63000 Clermont-Ferrand, France;	865
Email: <a href="mailto:franck.charmantray@uca.fr">franck.charmantray@uca.fr</a>	866
Bastien Doumèche – Univ Lyon, Université Claude Bernard	867
Lyon 1, 69622 Villeurbanne, France; <a href="https://orcid.org/0000-0003-3458-1029">orcid.org/0000-</a>	868
0003-3458-1029; Email: <a href="mailto:doumeche@univ-lyon1.fr">doumeche@univ-lyon1.fr</a>	869

Authors	870
Rainier-Numa Georges – Univ Lyon, Université Claude	871
Bernard Lyon 1, 69622 Villeurbanne, France	872
Lionel Ballut – Molecular Microbiology and Structural	873
Biochemistry, UMR 5086, CNRS-Université de Lyon, F-	874
69367 Lyon, France	875
Nushin Aghajari – Molecular Microbiology and Structural	876
Biochemistry, UMR 5086, CNRS-Université de Lyon, F-	877
69367 Lyon, France; <a href="https://orcid.org/0000-0002-2245-2679">orcid.org/0000-0002-2245-2679</a>	878
Laurence Hecquet – Université Clermont Auvergne, CNRS,	879
SIGMA Clermont, Institut de Chimie de Clermont-Ferrand	880
(ICCF), F-63000 Clermont-Ferrand, France; <a href="https://orcid.org/0000-0003-2971-5686">orcid.org/</a>	881
0000-0003-2971-5686	882

Complete contact information is available at: 883  
<https://pubs.acs.org/10.1021/acs.biochem.3c00714> 884

Author Contributions	885
The manuscript was written through contributions of all	886
authors. All authors have given approval to the final version of	887
the manuscript. R.-N.G. and L.B. contributed equally to this	888
work.	889

Funding	890
The Project CEITOP ( <i>Criblage Electrochimique d'Inhibiteur de</i>	891
<i>Transcétolases d'Organismes Pathogènes</i> ) is funded by the	892
Region Auvergne-Rhône-Alpes (AuRA).	893

Notes	894
The authors declare no competing financial interest.	895

■ ACKNOWLEDGMENTS	896
The AuRA region is greatly acknowledged for the funding of	897
one of us (R.-N.G.). The authors thank the SOLEIL	898
synchrotron for access to beamline PROXIMA-2A and the	899
European Synchrotron Radiation Facility for access to MX-	900
beamlines. Technical support from all the beamlines staff is	901
gratefully acknowledged. Louise Blanco-Garcia (Bachelor	902
student) is acknowledged for their help in the expression,	903
purification, and characterization of <i>ml</i> TK and <i>mf</i> TK.	904

## 905 ■ ABBREVIATIONS

906 *ab*: *Acinetobacter baumannii*; *cg*: *Candida glabrata*; *ec*:  
 907 *Escherichia coli*; *ef*: *Enterococcus faecium*; F6P: D-fructose-6-  
 908 phosphate;  $\text{Fe}(\text{CN})_6^{3-}$ : ferricyanide ion; *hi*: *Hemophilus*  
 909 *influenzae*; *hs*: *Homo sapiens*; *lm*: *Leishmania mexicana*; *ml*:  
 910 *Mycobacterium leprae*; *mt*: *Mycobacterium tuberculosis*; *pf*:  
 911 *Plasmodium falciparum*; *pa*: *Pseudomonas aeruginosa*; *sa*:  
 912 *Staphylococcus aureus*; *sp*: *Streptococcus pneumoniae*; TK:  
 913 transketolase; TPP: thiamine pyrophosphate; *vv*: *Vibrio*  
 914 *vulnificus*

## 915 ■ PROTEIN ID (UNIPROT)

916 *Acinetobacter baumannii* transketolase (*abTK*) B0VR28  
 917 *Arabidopsis thaliana* Q8RWV0  
 918 *Candida albicans* O94039  
 919 *Candida glabrata* transketolase (*cgTK*) Q6FWC3  
 920 *Columba livia*  
 921 A0A2I0M9P4 *Escherichia*  
 922 *coli* transke-  
 923 t o l a s e  
 924 (*ecTK*)  
 P27302  
 925 *Enterococcus faecium* transketolase (*efTK*) I3U1P7  
 926 *Geobacillus stearothermophilus*  
 927 A0A0I9QGZ2 *Hemophilus*  
 928 *influenzae*  
 929 transketo-  
 930 lase (*hiTK*)  
 P43757  
 931 *Homo sapiens* transketolase (*hsTK*) P29401  
 932 *Leishmania mexicana* transketolase (*lmTK*) Q8MPM3  
 933 *Maylandia zebra*  
 934 A0A3P9D0Z9 *Mycobacte-*  
 935 *rium leprae*  
 936 transketo-  
 937 lase (*mlTK*)  
 P46708  
 938 *Mycobacterium tuberculosis* transketolase (*mtTK*) P9WG25  
 939 *Oryza sativa* Q01LW1  
 940 *Pan paniscus*  
 941 A0A2R9AQ71 *Plasmodium*  
 942 *falciparum*  
 943 transketo-  
 944 lase (*pfTK*)  
 C6KSV3  
 945 *Pseudomonas aeruginosa* transketolase (*paTK*) Q9ISY8  
 946 *Saccharomyces cerevisiae* P23254  
 947 *Schistosoma japonicum* C1LGS7  
 948 *Schistosoma mansoni*  
 949 A0A3Q0KGI5 *S p i n a c i a*  
 950 *oleracea*  
 O20250  
 951 *Staphylococcus aureus* transketolase (*saTK*) P99161  
 952 *Streptococcus pneumoniae* transketolase (*spTK*) P22976  
 953 *Sus scrofa* A8U4R4  
 954 *Trypanosoma brucei gambiense* C9ZW96  
 955 *Vibrio cholerae* Q9KLW7  
 956 *Vibrio parahaemolyticus* Q87GY4  
 957 *Vibrio vulnificus* transketolase (*vvTK*) Q7MDD4  
 958 *Zea mays* Q7SIC9

959

## ■ REFERENCES

(1) Horecker, B.; Gibbs, M.; Klenow, H.; Smyrniotis, P. The 961  
 mechanism of pentose phosphate conversion to hexose mono- 962  
 phosphate. I. With a liver enzyme preparation. *J. Biol. Chem.* **1954**, 963  
 207 (1), 393–403. 964  
 (2) Gutowski, J.; Lienhard, G. Transition state analogs for thiamin 965  
 pyrophosphate-dependent enzymes. *J. Biol. Chem.* **1976**, 251 (9), 966  
 2863–2866. 967  
 (3) Kochetov, G.; Solovjeva, O. Structure and functioning 968  
 mechanism of transketolase. *Biochimica et Biophysica Acta (BBA) -* 969  
*Proteins and Proteomics* **2014**, 1844 (9), 1608–1618. 970  
 (4) Schneider, G.; Sundström, M.; Lindqvist, Y. Preliminary 971  
 crystallographic data for transketolase from yeast. *J. Biol. Chem.* 972  
**1989**, 264 (36), 21619–21620. 973  
 (5) Aymard, C.; Halma, M.; Comte, A.; Mousty, C.; Prévot, V.; 974  
 Hecquet, L.; Charmantray, F.; Blum, L.; Doumèche, B. Innovative 975  
 electrochemical screening allows transketolase inhibitors to be 976  
 identified. *Analytical chemistry* **2018**, 90 (15), 9241–9248. 977  
 (6) Sprenger, G.; Schörken, U.; Sprenger, G.; Sahm, H. Trans- 978  
 ketolase A of *Escherichia coli* K12. Purification and properties of the 979  
 enzyme from recombinant strains. *Eur. J. Biochem.* **1995**, 230 (2), 980  
 525–532. 981  
 (7) Tan, X.; Ramond, E.; Jamet, A.; Barnier, J.; Decaux-Tramoni, B.; 982  
 Dupuis, M.; Euphrasie, D.; Tros, F.; Nemazanyy, I.; Ziveri, J.; Nassif, 983  
 X.; Charbit, A.; Coureuil, M. Transketolase of *Staphylococcus aureus* 984  
 in the Control of Master Regulators of Stress Response During 985  
 Infection. *Journal of infectious diseases* **2019**, 220 (12), 1967–1976. 986  
 (8) Gabryszewski, S.; Wong Fok Lung, T.; Annavaajhala, M.; 987  
 Tomlinson, K.; Riquelme, S.; Khan, I.; Prince, A. Metabolic 988  
 Adaptation in Methicillin-Resistant *Staphylococcus aureus* Pneumonia. 989  
*Am. J. Respir. Cell Mol. Biol.* **2019**, 61 (2), 185–197. 990  
 (9) Cascante, M.; Centelles, J.; Veech, R.; Lee, W.; Boros, L. Role of 991  
 Thiamin (Vitamin B-1) and Transketolase in Tumor Cell 992  
 Proliferation. *Nutrition and Cancer* **2000**, 36 (2), 150–154. 993  
 (10) Boros, L.; Puigjaner, J.; Cascante, M.; Lee, W.; Brandes, J.; 994  
 Bassilian, S.; Schirmer, W. Oxythiamine and dehydroepiandrosterone 995  
 inhibit the nonoxidative synthesis of ribose and tumor cell 996  
 proliferation. *Cancer Res.* **1997**, 57 (19), 4242–4248. 997  
 (11) Bozdech, Z.; Ginsburg, H. Data mining of the transcriptome of 998  
*Plasmodium falciparum*: The pentose phosphate pathway and 999  
 ancillary processes. *Malar. J.* **2005**, 4 (1), 17. 1000  
 (12) Prejanò, M.; Medina, F.; Fernandes, P.; Russo, N.; Ramos, M.; 1001  
 Marino, T. The catalytic mechanism of Human transketolase. 1002  
*ChemPhysChem* **2019**, 20 (21), 2881–2886. 1003  
 (13) Fullam, E.; Pojer, F.; Bergfors, T.; Jones, T.; Cole, S. Structure 1004  
 and function of the transketolase from *Mycobacterium tuberculosis* and 1005  
 comparison with the human enzyme. *Open Biology* **2012**, 2 (1), 1006  
 No. 110026. 1007  
 (14) Mitschke, L.; Parthier, C.; Schröder-Tittmann, K.; Coy, J.; 1008  
 Lüdtke, S.; Tittmann, K. The crystal structure of Human transketolase 1009  
 and new insights into its mode of action. *J. Biol. Chem.* **2010**, 285 1010  
 (41), 31559–31570. 1011  
 (15) Lindqvist, Y.; Schneider, G.; Ermler, U.; Sundström, M. Three- 1012  
 dimensional structure of transketolase, a thiamine diphosphate 1013  
 dependent enzyme, at 2.5 Å resolution. *EMBO Journal* **1992**, 11 1014  
 (7), 2373–2379. 1015  
 (16) Asztalos, P.; Parthier, C.; Golbik, R.; Kleinschmidt, M.; 1016  
 Hübner, G.; Weiss, M.; Friedemann, R.; Wille, G.; Tittmann, K. Strain 1017  
 and near attack conformers in enzymic thiamin catalysis: X-ray 1018  
 crystallographic snapshots of bacterial transketolase in covalent 1019  
 complex with donor ketoses xylulose 5-phosphate and fructose 6- 1020  
 phosphate, and in noncovalent complex with acceptor aldose ribose 5- 1021  
 phosphate. *Biochemistry* **2007**, 46 (43), 12037–12052. 1022  
 (17) Schenk, G.; Duggleby, R.; Nixon, P. Properties and functions of 1023  
 the thiamin diphosphate dependent enzyme transketolase. *Internation-* 1024  
*al Journal of Biochemistry & Cell Biology* **1998**, 30 (12), 1297– 1025  
 1318. 1026

- 1027 (18) Fiedler, E.; Golbik, R.; Schneider, G.; Tittmann, K.; Neef, H.;  
1028 König, S.; Hübner, G. Examination of donor substrate conversion in  
1029 yeast transketolase. *J. Biol. Chem.* **2001**, *276* (19), 16051–16058.
- 1030 (19) Nauton, L.; Hélaïne, V.; Théry, V.; Hecquet, L. Insights into  
1031 the thiamine diphosphate enzyme activation mechanism: computa-  
1032 tional model for transketolase using a quantum mechanical/molecular  
1033 mechanical method. *Biochemistry* **2016**, *55* (14), 2144–2152.
- 1034 (20) Nauton, L.; Hecquet, L.; Théry, V. QM/MM study of Human  
1035 transketolase: thiamine diphosphate activation mechanism and  
1036 complete catalytic cycle. *J. Chem. Inf. Model.* **2021**, *61* (7), 3502–  
1037 3515.
- 1038 (21) Meshalkina, L.; Solovjeva, O.; Khodak, Y.; Druitsa, V.;  
1039 Kochetov, G. Isolation and properties of human transketolase.  
1040 *Biochemistry (Moscow)* **2010**, *75* (7), 873–880.
- 1041 (22) Meshalkina, L.; Solovjeva, O.; Kochetov, G. Interaction of  
1042 transketolase from Human tissues with substrates. *Biochemistry*  
1043 *(Moscow)* **2011**, *76* (9), 1061–1064.
- 1044 (23) Meshalkina, L.; Kochetov, G.; Brauer, J.; Hübner, G.; Tittmann,  
1045 K.; Golbik, R. New evidence for cofactor's amino group function in  
1046 thiamin catalysis by transketolase. *Biochemical and biophysical research*  
1047 *communications* **2008**, *366* (3), 692–697.
- 1048 (24) Fiedler, E.; Thorell, S.; Sandalova, T.; Golbik, R.; König, S.;  
1049 Schneider, G. Snapshot of a key intermediate in enzymatic thiamin  
1050 catalysis: Crystal structure of the  $\alpha$ -carbanion of ( $\alpha,\beta$ -dihydrox-  
1051 yethyl)-thiamin diphosphate in the active site of transketolase from  
1052 *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (2),  
1053 591–595.
- 1054 (25) Asokan, G.; Ramadhan, T.; Ahmed, E.; Sanad, H. WHO global  
1055 priority pathogens list: a bibliometric analysis of Medline-PubMed for  
1056 knowledge mobilization to infection prevention and control practices  
1057 in Bahrain. *Oman Med. J.* **2019**, *34* (3), 184–193.
- 1058 (26) [https://www.who.int/en/news/item/25-10-2022-who-releases-](https://www.who.int/en/news/item/25-10-2022-who-releases-first-ever-list-of-health-threatening-fungi)  
1059 [first-ever-list-of-health-threatening-fungi](https://www.who.int/en/news/item/25-10-2022-who-releases-first-ever-list-of-health-threatening-fungi), accessed on April 21, 2024.
- 1060 (27) [https://www.who.int/en/news-room/fact-sheets/detail/](https://www.who.int/en/news-room/fact-sheets/detail/leishmaniasis)  
1061 [leishmaniasis](https://www.who.int/en/news-room/fact-sheets/detail/leishmaniasis), accessed on April 21, 2024.
- 1062 (28) [https://www.who.int/en/news-room/fact-sheets/detail/](https://www.who.int/en/news-room/fact-sheets/detail/malaria)  
1063 [malaria](https://www.who.int/en/news-room/fact-sheets/detail/malaria), accessed on April 21, 2024.
- 1064 (29) Baker-Austin, C.; Oliver, J. *Vibrio vulnificus*: new insights into a  
1065 deadly opportunistic pathogen. *Environ. Microbiol.* **2018**, *20* (2), 423–  
1066 430. 2018 Feb;
- 1067 (30) Trinanés, J.; Martínez-Urtaza, J. Future scenarios of risk of  
1068 *Vibrio* infections in a warming planet: A global mapping study. *Lancet*  
1069 *Planetary Health* **2021**, *5* (7), e426–e435.
- 1070 (31) Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: an  
1071 online tool for phylogenetic tree display and annotation. *Nucleic Acids*  
1072 *Res.* **2021**, *49* (W1), W293–W296.
- 1073 (32) Datta, A.; Racker, E. Mechanism of action of transketolase. II.  
1074 The substrate-enzyme intermediate. *J. Biol. Chem.* **1961**, *236*, 624–  
1075 628.
- 1076 (33) De La Haba, G.; Leder, I.; Racker, E. Crystalline transketolase  
1077 from bakers' yeast: Isolation and properties. *J. Biol. Chem.* **1955**, *214*  
1078 (1), 409–426.
- 1079 (34) Kochetov, G.; Determination of transketolase activity via  
1080 ferricyanide reduction; In *Methods in Enzymology*, Elsevier, Vol. 89,  
1081 1982, pp. 43–44.
- 1082 (35) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello,  
1083 G.; Gumienny, R.; Schwede, T. SWISS-MODEL: Homology  
1084 modelling of protein structures and complexes. *Nucleic Acids Res.*  
1085 **2018**, *46* (W1), W296–W303.
- 1086 (36) Studer, G.; Rempfer, C.; Waterhouse, A.; Gumienny, R.; Haas,  
1087 J.; Schwede, T. QMEANDisCo-Distance constraints applied on  
1088 model quality estimation. *Bioinformatics* **2020**, *36* (6), 1765–1771.
- 1089 (37) Chen, V.; Arendall, W.; Headd, J.; Keedy, D.; Immormino, R.;  
1090 Kapral, G.; Richardson, D. MolProbity: all-atom structure validation  
1091 for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol.*  
1092 *Crystallogr.* **2010**, *66* (1), 12–21.
- 1093 (38) Guex, N.; Peitsch, M. SWISS-MODEL and the Swiss-Pdb  
1094 Viewer: an environment for comparative protein modeling. *Electro-*  
1095 *phoresis* **1997**, *18* (15), 2714–2723.
- (39) Bowie, J.; Lüthy, R.; Eisenberg, D. A method to identify protein  
sequences that fold into a known three-dimensional structure. *Science*  
**1991**, *253* (5016), 164–170.
- (40) Colovos, C.; Yeates, T. Verification of protein structures:  
patterns of nonbonded atomic interactions. *Protein Sci.* **1993**, *2* (9),  
1511–1519.
- (41) Laskowski, R.; MacArthur, M.; Moss, D.; Thornton, J.  
PROCHECK: a program to check the stereochemical quality of  
protein structures. *J. Appl. Crystallogr.* **1993**, *26* (2), 283–291.
- (42) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.;  
Ronneberger, O.; Hassabis, D. Highly accurate protein structure  
prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589.
- (43) Procter, J.; Carstairs, G.; Soares, B.; Mourão, K.; Ofoegbu, T.;  
Barton, D.; Barton, G. Alignment of Biological Sequences with  
Jalview. *Methods Mol. Biol.* **2021**, *2231*, 203–224.
- (44) Wilkins, M.; Gasteiger, E.; Bairoch, A.; Sanchez, J.; Williams,  
K.; Appel, R.; Hochstrasser, D. Protein identification and analysis  
tools in the ExPASy server. *Methods in molecular biology (Clifton, N.J.)*  
**1998**, *112*, 531–552.
- (45) Kabsch, W. XDS. *Acta Crystallographica Section D Biological*  
*Crystallography* **2010**, *66* (2), 125–132.
- (46) Afonine, P.; Mustyakimov, M.; Grosse-Kunstleve, R.; Moriarty,  
N.; Langan, P.; Adams, P. Joint X-ray and neutron refinement with  
phenix.refine. *Acta Crystallographica Section D Biological Crystallog-*  
*raphy* **2010**, *66* (11), 1153–1163.
- (47) Emsley, P.; Cowtan, K. Coot: Model-building tools for  
molecular graphics. *Acta Crystallographica Section D Biological*  
*Crystallography* **2004**, *60* (12), 2126–2132.
- (48) Liebschner, D.; Afonine, P.; Baker, M.; Bunkóczi, G.; Chen, V.;  
Croll, T.; Adams, P. Macromolecular structure determination using X-  
rays, neutrons and electrons: recent developments in Phenix. *Acta*  
*Crystallogr., Sect. D: Struct. Biol.* **2019**, *75* (10), 861–877.
- (49) Georges, R. N.; Ballut, L.; Octobre, G.; Comte, A.; Hecquet, L.;  
Charmantray, F.; Doumèche, B. Structural determination and kinetic  
analysis of the transketolase from *Vibrio vulnificus* reveal unexpected  
cooperative behaviour. *Protein Sci.* **2024**, *33* (3), No. e4884.
- (50) Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0:  
Computed atlas of surface topography of proteins. *Nucleic Acids Res.*  
**2018**, *46* (W1), W363–W367.
- (51) Clementel, D.; Del Conte, A.; Monzon, A.; Camagni, G.;  
Minervini, G.; Piovesan, D.; Tosatto, S. RING 3.0: fast generation of  
probabilistic residue interaction networks from structural ensembles.  
*Nucleic Acids Res.* **2022**, *50* (W1), W651–W656.
- (52) Martin, A.; Vidotto, M.; Boscaroli, F.; Di Domenico, T.; Walsh,  
I.; Tosatto, S. RING: networking interacting residues, evolutionary  
information and energetics in protein structures. *Bioinformatics* **2011**,  
*27* (14), 2003–2005.
- (53) Veitch, N.; Mauger, D.; Cazzulo, J.; Lindqvist, Y.; Barrett, M.  
Transketolase from *Leishmania mexicana* has a dual subcellular  
localization. *Biochem. J.* **2004**, *382* (2), 759–767.
- (54) Shcherbakova, T.; Baldin, S.; Shumkov, M.; Gushchina, I.;  
Nilov, D.; Švedas, V. Isolation and biochemical characterization of  
recombinant transketolase from *Mycobacterium tuberculosis*. *Acta*  
*Naturae* **2022**, *14* (2), 93–97.
- (55) Rabe Von Pappenheim, F.; Aldeghi, M.; Shome, B.; Begley, T.;  
De Groot, B. L.; Tittmann, K. Structural basis for antibiotic action of  
the B1 antivitamin 2'-methoxy-thiamine. *Nat. Chem. Biol.* **2020**, *16*  
(11), 1237–1245.
- (56) Wood, T.; Fletcher, S. The affinity chromatography of  
transketolase. *Biochim. Biophys. Acta* **1978**, *527* (1), 249–255.
- (57) Robinson, B.; Chun, K. The relationships between trans-  
ketolase, yeast pyruvate decarboxylase and pyruvate dehydrogenase of  
the pyruvate dehydrogenase complex. *FEBS letters* **1993**, *328* (1–2),  
99–102.
- (58) Maslova, A.; Meshalkina, L.; Kochetov, G. Computer modeling  
of transketolase-like protein, TKTL1, a marker of certain tumor  
tissues. *Biochemistry (Moscow)* **2012**, *77* (3), 296–299.
- (59) Nilsson, U.; Meshalkina, L.; Lindqvist, Y.; Schneider, G.  
Examination of substrate binding in thiamin diphosphate- dependent

1165 transketolase by protein crystallography and site-directed muta-  
1166 genesis. *J. Biol. Chem.* **1997**, *272* (3), 1864–1869.

1167 (60) Hasan, Md. A.; Mazumder, Md. H.; Chowdhury, A.; Datta, A.;  
1168 Khan, Md. A. Molecular-docking study of malaria drug target enzyme  
1169 transketolase in *Plasmodium falciparum* 3D7 portends the novel  
1170 approach to its treatment. *Source Code Biol. Med.* **2015**, *10* (1), 7.

1171 (61) Joshi, S.; Singh, A.; Kumar, A.; Misra, P.; Siddiqi, M.; Saxena, J.  
1172 Molecular cloning and characterization of *Plasmodium falciparum*  
1173 transketolase. *Mol. Biochem. Parasitol.* **2008**, *160* (1), 32–41.

1174 (62) König, S.; Schellenberger, A.; Neef, H.; Schneider, G.  
1175 Specificity of coenzyme binding in thiamin diphosphate-dependent  
1176 enzymes. Crystal structures of yeast transketolase in complex with  
1177 analogs of thiamin diphosphate. *J. Biol. Chem.* **1994**, *269* (14),  
1178 10879–10882.

1179 (63) Galperin, M.; Walker, D.; Koonin, E. Analogous enzymes:  
1180 independent inventions in enzyme evolution. *Genome research* **1998**, *8*  
1181 (8), 779–790.

1182 (64) Stevens, F. Homology versus analogy: possible evolutionary  
1183 relationship of immunoglobulins, cupredoxins, and Cu,Zn-superoxide  
1184 dismutase. *Journal of molecular recognition: JMR* **2008**, *21* (1), 20–29.

1185 (65) Zerilli, M.; Amato, M.; Martorana, A.; Cabibi, D.; Coy, J.;  
1186 Cappello, F.; Rodolico, V. Increased expression of transketolase-like-1  
1187 in papillary thyroid carcinomas smaller than 1.5 cm in diameter is  
1188 associated with lymph-node metastases. *Cancer* **2008**, *113* (5), 936–  
1189 944.