Structural comparisons of phosphoenolpyruvate carboxykinases reveal the evolutionary trajectories of these phosphodiester energy conversion enzymes

Received for publication, September 2, 2019; in revised form, October 24, 2019. Published, Papers in Press, October 28, 2019, DOI 10.1074/jbc.RA119.010920

Yoko Chiba (千葉洋子)1,2, Takuya Miyakawa (宮川拓也)3¹, Yasuhiro Shimane (嶋根康弘)4, Ken Takai (高井研)4, Masaru Tanokura (田之谷優)⁴, and Tomoyoshi Nozaki (野崎智義)⁴

From the ¹Department of Subsurface Geobiological Analysis and Research, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15, Natsushima-cho, Yokosuka-city, Kanagawa, 237-0061, Japan, the ²Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo, 113-8657, Japan, the ³Super-Cutting-Edge Grand and Advanced Research Program, Institute for Extra-Cutting-Edge Science and Technology Avant-Garde, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15, Natsushima-cho, Yokosuka-city, Kanagawa, 237-0061, Japan, and the ⁴Department of Biomedical Chemistry, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Edited by Joseph M. Jez

Inorganic pyrophosphate (PPi) consists of two phosphate molecules and can act as an energy and phosphate donor in cellular reactions, similar to ATP. Several kinases use PPi as a substrate, and these kinases have recently been suggested to have evolved from ATP-dependent functional homologs, which have significant amino acid sequence similarity to PPi-utilizing enzymes. In contrast, phosphoenolpyruvate carboxykinase (PEPCK) can be divided into three types according to the phosphate donor (ATP, GTP, or PPi), and the amino acid sequence similarity of these PEPCKs is too low to confirm that they share a common ancestor. Here we solved the crystal structure of a PPi-PEPCK homolog from the bacterium Actinomyces israelii at 2.6 Å resolution and compared it with previously reported structures from ATP- and GTP-specific PEPCKs to assess the degrees of similarities and divergences among these PEPCKs. These comparisons revealed that they share a tertiary structure with significant value and that amino acid residues directly contributing to substrate recognition, except for those that recognize purine moieties, are conserved. Furthermore, the order of secondary structural elements between PPi-, ATP-, and GTP-specific PEPCKs was strictly conserved. The structure-based comparisons of the three PEPCK types provide key insights into the structural basis of PPi, specificity and suggest that all of these PEPCKs are derived from a common ancestor.

Inorganic pyrophosphate (PPi)³ is the simplest compound containing a high-energy phosphate bond (1) between two Pi molecules. PPi can act as an energy and phosphate donor, similar to nucleoside di- or triphosphates, including ATP, in cellular reactions (2–4). Several enzymes selectively utilize PPi over ATP and other nucleotides to catalyze similar reactions as ATP-dependent functional homologs.

PPi-utilizing enzymes are potentially useful for metabolic engineering. In contrast to ATP, reactions involving PPi are reversible in vivo, with only a few exceptions (5–7) because the energy released by cleavage of PPi is smaller than that of ATP or GTP (4, 8). In addition, utilization of PPi requires less cellular energy because PPi is generated as a byproduct of many in vivo reactions hydrolyzing nucleotide triphosphate (4), whereas energy is required for ATP synthesis. A number of organisms, particularly anaerobic fermenting microbes, utilize PPi-dependent enzymes instead of ATP-dependent functional homologs for glycolysis and closely related metabolic reactions as a strategy to increase net ATP production (9). Therefore, introducing or substituting PPi-dependent enzymes in place of ATP-dependent functional homologs for glycolysis and closely related metabolic reactions is an attractive approach to alter metabolic flux or improve cellular energy efficiency. Furthermore, use of PPi-dependent enzymes is an attractive possibility for industrial production of phosphorylated compounds because PPi is 1000 times cheaper than ATP (10).

The functional and structural study of PPi-utilizing enzymes is expected to provide important insights into the evolution of cellular energy currency. Because of its simple structure, PPi has been proposed to be the evolutionary precursor of ATP (3, 11).

This work was supported by the Platform Project for Supporting Drug Discovery and Life Science Research (Platform for Drug Discovery, Informatics, and Structural Life Science) from the Japan Agency for Medical Research and Development (to M. T.), and the Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research from the Japan Agency for Medical Research and Development under Grant JP19am0101077) (to T. M.). This work was partially supported by Japan Society for the Promotion of Science KAKENHI Grants 26860275 and 17H05242 (to Y. C.). The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Figs. S1–S8.

The atomic coordinates and structure factors (code 6K3I) have been deposited in the Protein Data Bank (http://wwpdb.org/).

1 Both authors contributed equally to this work.
2 To whom correspondence should be addressed: RIKEN Center for Sustainable Resource Science, 2-1 Hirosawa, Wako-shi, Saitama, 351-0198, Japan. Tel.: 81-48-467-9372; E-mail: ychiba@jamstec.go.jp.

³The abbreviations used are: PPi, inorganic pyrophosphate; PFK, phosphofructokinase; ACK, acetate kinase; PEPCK, phosphoenolpyruvate carboxykinase; Ai, Actinomycetes israelii; Pf, Propionibacterium freudenreichii subsp. shermanii; Eh, Entamoeba histolytica; RMSD, root mean square deviation; FHA, forkhead-associated; LUCA, last universal common ancestor.
In addition, adoption of PP<sub>i</sub>-using enzymes by ancestral organisms with poor ATP-producing ability may have been energetically favorable. Therefore, understanding the evolutionary origin and relationship of PP<sub>i</sub>-dependent enzymes with ATP-dependent functional homologs is of interest for the evolutionary study of metabolism.

The evolutionary relationship between PP<sub>i</sub>-dependent kinases and ATP-dependent functional homologs has been discussed in several studies. The best-studied enzyme to date is acetate kinases (ACKs) utilize ATP (EC 2.7.2.1) as the phosphate donor. The only exception is an ACK from the eukaryotic parasite *Entamoeba histolytica* that strictly recognizes PP<sub>i</sub> as the substrate (EC 2.7.2.12) (7, 17). PP<sub>i</sub>-ACK has clear homology to ATP-ACKs, and phylogenetic analysis indicated that PP<sub>i</sub>-ACK has not formed a separate clade from ATP-ACKs (18, 19), suggesting that PP<sub>i</sub>-ACK also arose from an ATP-ACK. In addition, a member of the ribokinase family of proteins, which are considered to be ATP- or ADP-dependent, was recently found to be PP<sub>i</sub>-dependent (10). In summary, these three PP<sub>i</sub>-dependent kinases and their functional ATP-utilizing homologs have clear amino acid sequence similarity, and the former appear to have evolved from ATP-dependent ancestors.

The evolution of phosphoenolpyruvate carboxykinase (PEPCK) seems to have followed a different pathway from the kinases described above. Depending on the phosphate donor to oxaloacetate, PEPCK can be divided into three types: GTP-PEPCK (EC 4.1.1.32), ATP-PEPCK (EC 4.1.1.49), and PP<sub>i</sub>-PEPCK (EC 4.1.1.38). Although both ATP- and GTP-PEPCK show significant amino acid sequence identity within each type, no significant overall sequence homology is observed between the two groups of enzymes (20). In contrast, crystal structure-based studies have revealed that ATP- and GTP-PEPCK have highly similar tertiary structures (21, 22). These enzymes possess “consensus motifs,” including a PEPCK-specific domain, which directly associates with PEP or oxaloacetate, and a P-binding loop (also called kinase-1a) and kinase-2 motifs, which directly interact with the phosphate moiety of nucleoside triphosphates (20, 21). In contrast to ATP- and GTP-PEPCK, PP<sub>i</sub>-PEPCK has not been structurally characterized at the tertiary level. Although a PP<sub>i</sub>-PEPCK was biochemically characterized more than 50 years ago (23–27), the amino acid and gene sequence of the enzyme has recently been reported (28). PP<sub>i</sub>-PEPCK consists of more than 1100 amino acid residues and is approximately twice as long as ATP- and GTP PEPCK. PP<sub>i</sub>-PEPCK does not share a significant amino acid sequence identity with ATP- or GTP-PEPCK (E-value $\geq$ 1). It remains to be determined whether PP<sub>i</sub>-PEPCK has a similar tertiary structure as ATP- and GTP-PEPCK.

### Table 1

| Data collection and refinement statistics |
|-----------------------------------------|
| Beamline                                 | PF BL-17A |
| Wavelength (Å)                          | 0.97887 |
| Space group                             | P3<sub>2</sub>21 |
| Cell dimensions: a, b, c (Å)            | 160.4, 160.4, 200.2 |
| Resolution (Å)                          | 48.1–2.60 |
| No. of unique reflections               | 91847 (450S) |
| R<sub>work</sub>                         | 0.158 (1.243) |
| R<sub>free</sub>                         | 0.049 (0.389) |
| CC[ifrax,1/2]                           | 0.999 (0.837) |
| Mean I/σ(I)                             | 19.2 (3.3) |
| Completeness (%)                        | 99.9 (83.7) |
| Multiplicity                            | 20.1 (19.8) |

* Values in parentheses indicate those of the highest-resolution shell.

In this study, we obtained and compared the crystal structure of a PP<sub>i</sub>-PEPCK homolog from *Actinomyces israelii* (AiPEPCK) with those of ATP- and GTP-PEPCK to evaluate the degree of homology between PP<sub>i</sub>-, ATP-, and GTP-PEPCK and to determine the structural basis of the PP<sub>i</sub> specificity of PP<sub>i</sub>-PEPCK.

### Results and discussion

**PP<sub>i</sub>-dependent activity and structure determination of PEPCK**

Crystallization was attempted using the biochemically characterized PP<sub>i</sub>-PEPCKs from *Propionibacterium freudenreichii* subsp. *shermanii* (PiPEPCK, WP_013160152.1) and *E. histolytica* (EhPEPCK1, XP_654765.1) (28) and PP<sub>i</sub>-PEPCK homologs from several bacteria. The crystals of PiPEPCK and the PP<sub>i</sub>-PEPCK homolog from *A. israelii* (AiPEPCK, WP_043560275.1), which show 61% and 43% amino acid identity with PiPEPCK and EhPEPCK1, respectively, were obtained and subjected to X-ray analysis. Quality diffraction data were collected only from a crystal of AiPEPCK labeled with selenomethionine. The data collection and refinement statistics are presented in Table 1.

PP<sub>i</sub>-dependent PEPCK activity was detected ($2 \times 10^{-2}$ μmol·mg of protein$^{-1}$·min$^{-1}$) from purified AiPEPCK, whereas ATP- and GTP-dependent activity was not detected. The activity was not increased when Mn was added to the reaction mixture instead of Co. The PP<sub>i</sub>-PEPCK activity was two orders of magnitude lower than that of PiPEPCK (28) under the same reaction conditions. The lower PP<sub>i</sub>-PEPCK activity of AiPEPCK may have been attributable to the purification and experimental conditions; for example, AiPEPCK might have lower oxygen tolerance than PiPEPCK, resulting in reduced activity following purification under aerobic conditions.
Because AiPEPCK has high (61%) amino acid sequence identity to biochemically characterized PPiPEPCK, and strictly conserved residues are likely critical for the PEPCK activity of PPiPEPCK and EhPEPCK, as described above, we considered that at least the three-dimensional structure of monomeric AiPEPCK is highly similar to that of the PPiPEPCKs from \textit{P. freudenreichii} and \textit{E. histolytica}. The structure of AiPEPCK was therefore used as the representative structure of PPiPEPCKs.

### Overall structure of AiPEPCK

The crystal structure of AiPEPCK was determined at 2.6 Å resolution. Two molecules in an asymmetric unit contained 1117 amino acid residues (A chain: 8–28, 39–431, 435–523, 542–755, and 758–1149) and 1118 residues (B chain: 11–28, 39–431, 435–523, and 542–1149) (Fig. S1). A Co ion was observed in each molecule, a finding that was not unexpected, as AiPEPCK was crystallized in the presence of Co and Mg ions, which are both required for EhPEPCK activity (26). The crystal structure of AiPEPCK was highly similar between the A and B chains (Cα RMSD of 0.141 Å); therefore, the AiPEPCK structure was described using the A chain, which had a lower B factor (49.0 Å²) than the B chain (61.4 Å²) (Fig. S1).

The AiPEPCK structure is divided into a core structure that grasps a Co ion and four lobe structures (lobes 1 to 4) that envelope the core structure (Fig. 1, A and B). The core structure is composed of two globular α/β domains, called the N-terminal (85–346, 685–721, and 758–792) and C-terminal domains (347–359, 475–496, 649–684, and 817–1047).

N-terminal domain adopts a six-stranded β-sheet (S1, the order of strands is 1-2-3-31-32-30) and a seven-stranded β-sheet (S2, the order of strands is 6-7-8-10-9-4-5), and three α-helices (α4, α7, and α10) are sandwiched between these two β-sheets. Lobe 1 (1–84) exists in the N terminus of the N-terminal domain and is composed of three α-helices (α1, α2, and α3) and a 3_{10}-helix (η1). Helix α3 contacts S2 in the N-terminal domain. S2 also contacts four α-helices (α5, α6, α8, and α9) and two 3_{10}-helices (η2 and η3). Lobe 3 (722–757), which is inserted between β31 and β32 of the N-terminal domain, contains three α-helices (α17–α19). Lobe 3 is located on the opposite face of S1 from S2 together with a helix α16.

The C-terminal domain has a twisted β-sheet composed of 10 β-strands (S3, the order of strands is 29-28/27-11-23-33-35-34-39-36) and a β-hairpin that is inserted between β36 and β39 (Fig. 1A). S3 is surrounded by seven α-helices (α22–α28) and four 3_{10}-helices (η4, η8, and η9). The C-terminal domain contacts lobe 2 (360–474, 497–648, and 793–816) and lobe 4 (1048–1149) in addition to the N-terminal domain. Lobe 2 seems to adopt a large domain with an α + β structure, unlike other smaller lobes, and is divided into three parts in the primary structure (Fig. 1B). The N-terminal part of lobe 2 is inserted between β11 and β23 of the C-terminal domain and forms a β-sandwich fold with two antiparallel β-sheets (S4, the order of strands is 12-22-21-20-16-15-14, where β16 and β20 are parallel; S5, the order of strands is 13-17-18-19) and a helix α11. S4 contacts a part of helix α21 (793–816) and a portion of lobe 2 and connects α20 of the N-terminal domain and β33 of...
the C-terminal domain. The other part of lobe 2 (497–648) is inserted between H9257 and H9252 of the C-terminal domain and forms four H9251-helices (H925112–H925115), three 310-helices (H92575–H92577), and a H9252-hairpin (H925225 and H925226) with an attached H9252-strand (H925224). On the other hand, lobe 4 is extended from H925128 of the C-terminal domain and adopts a helical structure composed of six H9251-helices (H925129–H925134) and a 310-helix (H925713). Helices H925129 and H925130 contact helices H925122, H925123, and H925128 of the C-terminal domain. Helices H925131–H925134 cover the loops connecting the H9252-strands of S1 in the N-terminal domain.

**Dimer formation with the contacts between lobes 2 and 3**

To evaluate the oligomeric state of AiPEPCK, we carried out size-exclusion chromatography. The result indicated that soluble AiPEPCK exists mainly as a homodimer (Fig. 2A). Dimer formation was further analyzed using the PISA server (29), and each chain was predicted to form the same homodimer with an identical chain generated by symmetry operation (Fig. 2B). The contact surface area between two protomers was 1605 Å² (A-A' dimer) and 1599 (B-B' dimer).

In the quaternary structure of AiPEPCK, lobes 2 and 3 are located on the dimer interface (Fig. 2B). These lobes seem to form handclasp-like interactions; lobe 2 contacts lobes 2’ and 3’ of another protomer, and lobe 3 is positioned in close proximity to lobe 2’. The strand β26 of lobe 2 forms three main-chain hydrogen bonds, which are observed in a parallel β-sheet, with the loop connecting η7 and β25’ in lobe 2’ (Fig. 2C). This dimer interface is further reinforced by two hydrogen bonds between Ser640 and Asn607 and a van der Waals contact between Trp636 and Gly618. On the other hand, van der Waals contacts mainly contribute to the dimer interface between lobes 3 and 2’ (Fig. 2D), and Asp731 and Ser736 in lobe 3 also form hydrogen bonds with Gln425 and Trp467 in lobe 2’. These residues located on α17 and α18 of lobe 3 and on the loops between Ser414 and Ser431 in lobe 2’. There is no interaction among other lobes and the core structure (Fig. 2B). These structural findings suggest that lobes 2 and 3 are required for dimer formation of AiPEPCK.

**Structure comparison of PPi-PEPCK with ATP- and GTP-PEPCK**

A structural similarity search using the DALI server (30) revealed that PPi-PEPCK has significant similarity (Z score ≥ 10) only to ATP- and GTP-PEPCK. The top hit was ATP-PEPCK from Escherichia coli (PDB code 1OS1-A; Z score, 21.6;
Structural comparisons of phosphoenolpyruvate carboxykinases

RMSD, 3.6 Å; sequence identity, 13%), and the top hit among GTP-PEPCK was an enzyme from Rattus norvegicus (PDB code 5FH0-A; Z score, 17.5; RMSD, 3.5 Å; sequence identity, 10%). The structural superposition of PP1\textsubscript{r}, ATP-, and GTP-PEPCK revealed that they share a core structure consisting of the N- and C-terminal domains (Fig. S2), although PP1\textsubscript{r} is~500 amino acid residues longer than ATP- and GTP-PEPCK by lobe structure (lobes 1–4, Figs. 2A). Notably, the order of the secondary structural elements present in the core structure of PP1\textsubscript{r}, ATP-, and GTP-PEPCK was completely conserved in the primary structures of all enzymes (Fig. S3). A Co ion was located in the deep cleft between the two globular domains of PP1\textsubscript{r}. ATP- and GTP-PEPCK adopt Ca and Mn ions at the same position, respectively (Fig. S2). In addition, the residues surrounding the metal ions (Lys\textsubscript{333}, Lys\textsubscript{332}, His\textsubscript{352}, Asp\textsubscript{655}, and Asp\textsubscript{656} in AiPEPCK) are spatially conserved among three types of PEPCKs. The Mn ion at this position is required for enzymatic activity of ATP- and GTP-PEPCK (22, 31), indicating that the cleft likely functions as the active site in PP1\textsubscript{r}.

The lobe structures are specific to PP1\textsubscript{r} and are not contained in ATP- and GTP-PEPCKs. Amino acid sequence comparisons revealed that lobes 1–4 existed in all PP1\textsubscript{r}s (Fig. S4). We further searched for reported structures similar to individual lobe structures using the DALI server by extracting them from the overall structure of AiPEPCK. As a result, there was no protein hit with query of lobe 1, whereas the overall structure of lobe 3 and the partial structures of lobes 2 and 4 matched structural elements of other proteins.

The DALI results indicated that the overall structural similarity with lobe 2 was not found in any other proteins. However, the structure of lobe 2 was partially related to those of various types of proteins, such as a component of the bacterial type VII secretion apparatus EssC, a putative transcriptional regulator of the arabinosyltransferase EmbR, propionyl-CoA synthetase, the serine phosphate MtX, adenylate cyclase-like protein CT664, and so on (Fig. S5). Among the structural elements of lobe 2, the β-sandwich fold of β-sheets S4 and S5 highly emerges in the protein structures found in the DALI analysis. The top hit was EssC from Staphylococcus aureus (PDB code 1WV3-A; Z score, 4.3; RMSD, 2.9 Å; sequence identity, 6%). This protein and CT664 have a forkhead-associated (FHA) domain that adopts a β-sandwich fold and functions as a phosphopeptide recognition module (32). Although the loops connecting β-sheets S4 and S5 are used for peptide recognition of lobe 3, similar to the FHA domain (Fig. 2D), lobe 2 shows no sequence similarity to the FHA domain, and lobe 3 has no phosphorylated residue. The structure of lobe 3 was similar to the structural element of NUP155 (1166–1195), a component of the human nuclear pore complex (PDB code 5JJN-E and 5JJO-E; Z score, 2.1; RMSD, 2.4 Å; sequence identity, 6%) (Fig. S6) and a large protein with 1391 residues. According to the structures determined by cryoelectron tomography (33), this structural element is located outside of the pore ring and seems not to interact with any other subunits. In addition, the residues on the dimer interface of lobe 3 are not conserved in the similar structural element of NUP155. The helices of lobe 2 (497–648) also emerge as a partial structure of the large globular protein, as shown in propionyl-CoA synthetase (Fig. S5).

However, the structural elements of lobe 2–2’ interaction (Fig. 2C) are not conserved in any proteins, according to DALI results. These structural findings suggest that lobes 2 and 3 may be specific modules for dimer formation of PP1\textsubscript{r}.

Structures similar to a part of lobe 4 were found in various types of proteins with antiparallel helices consisting of two long α-helices (Fig. S7). Five protein structures with the highest Z scores are as follows: two molecules in the Rad50 dimer, a component of the Mre11 complex for the eukaryotic DNA damage response; PHYL1, a phyllody-inducing effector protein of phytoplasma; seryl-tRNA synthetase; and Sso2, a t-SNARE (target-membrane-associated–soluble N-ethylmaleimide fusion protein attachment protein SNAP receptor) protein that functions in intracellular membrane fusion. The top hit was one molecule in the Rad50 dimer (PDB code 1GOX-A; Z score, 6.2; RMSD, 5.1 Å; sequence identity, 6%). Their antiparallel helices match helices α31 and α33 of lobe 4, whereas their functions are highly divergent. Rad50 uses the antiparallel helices as an arm for assembly of the Mre11 complex (34). This structural element is required for tetramer formation and folding of a four-helix bundle in PHT1 and Sso2, respectively (35, 36). On the other hand, seryl-tRNA synthetase interacts with tRNA using parallel helices (37). In AiPEPCK, helices α31 and α33 of lobe 4 contact the loops connecting the β-strands of S1 in the N-terminal domain. Therefore, the structural homology search could not identify the functional role and evolutionary origin of lobe 4.

Conservation of residues important for PEPCK activity

Standard multiple sequence alignment algorithms work well for amino acid sequences with high similarity but are not suitable for sequences with low similarity or largely different lengths. Here a structure-based alignment algorithm was employed to perform an amino acid sequence alignment of PP1\textsubscript{r}, ATP-, and GTP-PEPCK because they share less than 10% amino acid sequence similarity, and PP1\textsubscript{r} is twice as long as ATP- and GTP-PEPCK. The structure-based alignment revealed that the catalytic residues in ATP- and GTP-PEPCK are strictly conserved in AiPEPCK and other PP1\textsubscript{r}s (Figs. 3A and Fig. S3). Most of the conserved residues are located within three motifs found in ATP- and GTP-PEPCK: a PEPCK-specific domain, P-binding loop, and kinase 2 domain, which interact with PEP or oxaloacetate, nucleotide triphosphate, and a divergent cation, respectively (20, 38). This finding strongly suggests that PP1\textsubscript{r} also contains these three motifs and utilizes a similar catalytic mechanism.

All of the conserved catalytic residues (Arg\textsuperscript{477}, Tyr\textsuperscript{435}, Gly\textsuperscript{327}, Lys\textsuperscript{334}, Ser\textsuperscript{368}, Asn\textsuperscript{393}, and Arg\textsuperscript{395}; rat cytosolic GTP-PEPCK numbering) in ATP and/or ATP-PEPCK that directly interact with PEP, oxaloacetate, or the analog oxalate (20, 22, 39, 40) were conserved in AiPEPCK as Arg\textsuperscript{334}, Tyr\textsuperscript{352}, Gly\textsuperscript{332}, Lys\textsuperscript{333}, Ser\textsuperscript{388}, and Arg\textsuperscript{371}, respectively (Fig. 3A, red triangles). The amino acid residues making hydrogen bonds to the triphosphate group of ATP in ATP-PEPCK (His\textsuperscript{232}, Ser\textsuperscript{250}, Gly\textsuperscript{251}, Gly\textsuperscript{253}, Lys\textsuperscript{254}, Thr\textsuperscript{255}, and Arg\textsuperscript{333}) (41) were also conserved in PP1\textsubscript{r} because of His\textsuperscript{233}, Ser\textsuperscript{249}, Gly\textsuperscript{246}, Gly\textsuperscript{248}, Lys\textsuperscript{250}, Ser\textsuperscript{252}, and Arg\textsuperscript{371}, respectively (Fig. 3A, orange circles). The
 orientation of the side chains was also similar between ATP/ GTP-PEPCK and PPi-PEPCK. In addition, all residues ligand to the Co ion (Lys332, His352, and Asp656) in PPi-PEPCK (Fig. 3A, pink circles) were conserved in ATP/GTP-PEPCK.

**Structural basis of PPi specificity**

PP1-PEPCK does not recognize ADP or GDP as a phosphate acceptor (24). As expected, the amino acid residues in purine-binding regions were not conserved between ATP-PEPCK (449RISIKDT455) (20), GTP-PEPCK (516WFRKDKNGKFLWPGFGENSRVRLEWMF) (22, 42), and PPi-PEPCK (Fig. 3A, pink circles) were conserved in ATP/GTP-PEPCK.

**Figure 3. Comparison of active sites.** A, comparison of residues in the active sites of PPi-, ATP-, and GTP-PEPCK. Red arrows, residues that interact with PEP, oxaloacetate, or analogs; orange circles, residues that interact with the phosphogroup of ATP and/or GTP; pink circles, residues that interact with a divalent cation. His920 of AiPEPCK, which is suggested to contribute to PPi specificity, is highlighted by a yellow box. B and C, comparison of the active site structure between AiPEPCK (B, light blue) and ATP-binding ATP-PEPCK (PDB code 2PXZ, yellow) and AiPEPCK and GTP-binding GTP-PEPCK (C, PDB code 3DT7, light brown). Residues of PPi-PEPCK and ATP- or GTP-PEPCK are indicated by regular letters and letters in parentheses, respectively.

PPi-PEPCK shows structural comparisons of phosphoenolpyruvate carboxykinases
Structural comparisons of phosphoenolpyruvate carboxykinases

Figure 4. A and B, comparison of the purine-recognizing helix and P-loop positions between ATP-PEPCK with (PDB code 2PXZ, yellow) and without (PDB code 10EN, light green) ATP (A) and AiPEPCK (light blue) and ATP-PEPCK without ATP (B). The purine-recognizing helix of AiPEPCK is closed, as in the case of ATP-binding ATP-PEPCK, whereas the P-loop of PPi-PEPCK is open, as in the case of apoATP-PEPCK.

Figure 5. Distribution of the three types of PEPCKs in extant organisms and the predicted evolutionary history.

Evolutionary relationship between three types of PEPCKs

ATP-PEPCK and GTP-PEPCK share less than 20% overall amino acid sequence identity and also share several conserved motifs, it has been considered that these two types of PEPCKs likely resulted from convergent evolution (20, 46). The amino acid sequence of PPi-PEPCK was also found to lack overall sequence similarity with ATP-PEPCK and GTP-PEPCK and is nearly as low as that of ATP-PEPCK and GTP-PEPCK (28), suggesting that PPi-PEPCK also does not share a common ancestor with ATP-PEPCK or GTP-PEPCK. Although an amino acid sequence–based search failed to detect conserved motifs in PPi-PEPCK, crystal structure analyses revealed that PPi-PEPCK conserves motifs common to ATP-PEPCK and GTP-PEPCK (Fig. 3A), and the three-dimensional structures of the three types of PEPCKs share statistically significant similarity. Furthermore, the order of the secondary structural elements was completely conserved in the primary structures, without exception (Fig. S3). These facts strongly suggest that PPi-, ATP-, and GTP-PEPCKs are not the result of convergent evolution but have a shared origin, as in the case of enzymes that share whole structures and the catalytic domains (47–50). If this assumption is true, then the timing of the divergence of the three types of PEPCKs and the function of the common ancestor remain to be determined.

The phylogenetic tree constructed using the structure-based alignment of core structures revealed that each type of PEPCK forms a single clade with significant bootstrap values (Fig. S8), indicating that the phosphate donor change did not occur multiple times, unlike in the case of PFK. It should be noted that the order of division cannot be discerned from this tree solely because outgroup sequences are not available. However, as PPi-PEPCK possesses insertion sequences forming appendage structures (lobes 1–4) that are absent in ATP-PEPCK and GTP-PEPCK, it is most probable that, at first, the ancestor of PPi-PEPCK diverged from the ancestor of three types of PEPCKs, and then the ancestor of ATP-PEPCK and GTP-PEPCK diverged (Fig. 5). Both ATP- and GTP-PEPCKs exist in various bacteria and archaea and follow a chimeric distribution. For example, most alpha-, gamma-, and epsilonproteobacteria and approximately half of all deltaproteobacteria have ATP-type PEPCKs, whereas nearly half of all betaproteobacteria and deltaproteobacteria possess GTP-type PEPCKs. Furthermore, in the case of eukaryotes, ATP-PEPCKs exist in yeasts and plants, whereas GTP-dependent PEPCKs are found in higher organisms, including animals and insects. Phylogenetic analyses indicate that most archaeal and bacterial ATP- and GTP-PEPCKs sequences are separated by significant values (Fig. S8). Taken together, these results suggest that the separation of ATP-PEPCK and GTP-PEPCK occurred before the birth of the last universal common ancestor (LUCA). To sum up these discussions, the ancestor of PPi-PEPCK may also separate from the common ancestor of the three types of PEPCKs before the LUCA. The three types of PEPCKs may have evolved independently for a sufficiently long period to lose amino acid sequence similarity but still retain the active site, with the exception of the purine-binding region.

It remains uncertain whether the common ancestor of the three types of PEPCKs possessed PPi-PEPCK specific lobes 1–4.
Structural comparisons of phosphoenolpyruvate carboxykinases

or whether the appendage structures were inserted after the ancestor of PP1-PEPCK diverged from the common ancestor of the three types of PEPCKs. If lobes 1–4 were added to ancestral PEPCK by lateral gene transfer, then sequences and/or structures that were the source of lobes 1–4 might remain in protein sequences or structures of extant organisms. However, BLASTP and PDB searches conducted using lobes 1–4 as queries found no sequences or functionally conserved structures like the core structure with statistically significant similarities. Discovery of enzymes that share the origin with PEPCKs is required to estimate the presence or absence of the appendage structures and substrate specificity of the common ancestor of PEPCKs.

The evolutionary history of the three types of PEPCKs is clearly distinct from that of other kinases. In the case of PFK, the PP type has high amino acid sequence similarity to ATP types, and PP1-PEFCK appears to have been derived from ATP-PFK in multiple events that have occurred relatively recently and involved substitution of one or two Gly residues in the active site with bulky ones (13, 14). In contrast, separation of PP1-PEPCK and the ATP- and GTP-dependent functional homologs occurred only once, most probably before the LUCA arose. Furthermore, a drastic change of the structure (insertion or deletion of multiple appendage structures) occurred when the ancestor of ATP- and GTP-PEPCK and PP1-PEPCK were separated. In conclusion, the present structure-based analyses of PP1-PEPCK have helped determine the evolutionary history of PEPCKs, which could not have been detected from the amino acid sequence alone.

Materials and methods

Plasmid construction

AiPEPCK (WP_043560275.1) was PCR-amplified from genomic DNA (JGD12771) purchased from RIKEN BRC, which is participating in the National Bio-Resource Project of MEXT, Japan using the primers 5’-TCGAGGTAGGCCATA-ATGTCCGTAGTCGACGC-3’ and 5’-ATTCGGATCCCT-CGATCAGACGAACCTGGGCTG-3’. The amplified DNA was then cloned into pCold GST plasmids (Takara) cut with NdeI and XhoI.

Crystallization experiments were performed using commercially available crystallization kits (Crystal Screen HT, Index HT (Hampton Research, Aliso Viejo, CA), and Wizard Screens I and II (Emerald BioSystems, Bainbridge Island, WA) at 293 K in 96-well VIOLAMO sitting-drop protein crystallization plates (AS ONE Co., Osaka, Japan) with a Gryphon protein crystallization system (Art Robbins Instruments, Sunnyvale, CA). A sitting drop was prepared by mixing 0.2 μl of protein solution and 0.2 μl of reservoir solution and equilibrated against 40 μl of reservoir solution. After optimizing the crystallization conditions, the crystals obtained using a reservoir solution consisting of 100 mM HEPES-NaOH (pH 7.5), 30% (w/v) PEG 400, and 200 mM MgCl2 were used for data collection.

To determine the quaternary structure of AiPEPCK, gel filtration was performed using a Superdex 200 (10/300) column equilibrated with 20 mM Tris-HCl (pH 8.0) supplemented with 150 mM NaCl at a flow rate of 1 ml min⁻¹. Gel filtration standard (Bio-Rad, catalog no. 1511901) was used as a standard.

Crystallization, data collection, and preliminary X-ray analysis

The selenomethionine-modified AiPEPCK crystals were soaked in crystallization solution supplemented with 25% (v/v) ethylene glycol as a cryoprotectant, picked up using Dual Thickness MicroMounts™ (MiTeGen, Ithaca, NY), and cooled in a liquid nitrogen stream. The X-ray diffraction data (1800 images) were collected using a PILATUS 6 M detector on beamline BL-17A at the Photon Factory (Tsukuba, Japan) with the following parameters: wavelength, 0.97887 Å (Se peak); oscillation angle, 0.2°; exposure time, 0.5 s; crystal-to-detector distance, 485.7 mm.
The diffraction data were indexed, integrated, and scaled using XDS (51) and AIMLESS (52). The obtained crystal belonged to the space group P31,21 with unit cell parameters of \( a = b = 160.4 \text{ Å} \) and \( c = 200.2 \text{ Å} \). The initial model was solved by single-wavelength anomalous dispersion phasing using AutoSol of the PHENIX program suite (53). Iterative model building and refinement cycles were performed using COOT (54) and PHENIX.REFINE (55). The final model was refined by Refmac5 (56) with twin refinement (twin fraction of 0.062) and local noncrystallographic symmetry restraint to an \( R_{\text{work}} \) of 0.196 and \( R_{\text{free}} \) of 0.241. Data collection and refinement statistics are summarized in Table 1. All structures were depicted using PyMOL viewer (Schrödinger, Tokyo, Japan).

**Construction of structure-based amino acid sequence alignment**

6K31 (A. israelii PP-PEPCK), 2PXZ (E. coli ATP-PEPCK), and 3DT7 (rat cytosolic GTP-PEPCK) (57) were used for structure-based amino acid sequence alignment. Structures with complete sequences were reconstructed as follows because the PDB coordinate files lacked some amino acid residues. Complete sequences of each PDB structure were obtained from GenBank. Self-homology modeling with each complete sequence was performed using SWISS-MODEL (58). The first through seventeenth and first through eighth sequences of AliPEPCK and ATP-PEPCK, respectively, were added manually because the program was unable to determine the topology.

The structure-based alignment of the reconstructed and full-length PEPCks was performed using the MultiSeq plugin (59) in the VMD program (60) and visualized with ESPript 3.0 (61). Secondary structures were predicted using PyMOL.

**Author contributions**—Y. C. conceptualization; Y. C. data curation; Y. C., T. M., and Y. S. formal analysis; Y. C. methodology; Y. C. and T. M. writing-original draft; Y. C., K. T., and M. T. writing-review and editing.

**Acknowledgments**—We thank Mihoko Imada (National Institute of Infectious Diseases) and Keiko Usui (Japan Agency for Marine-Earth Science and Technology; IAMSTEC) for plasmid construction. The synchrotron radiation experiments were performed on beamline BL-17A at the Photon Factory with approval from the High Energy Accelerator Research Organization (proposal 2016G648).

**References**

1. Ma, B., Meredith, C., and Schaefer III, H. F. (1994) Pyrophosphate structures and reactions: evaluation of electrostatic effects on the pyrophosphates with and without alkali cations. *J. Phys. Chem.* 98, 8216–8223 CrossRef

2. Baltschekfisky, M., and Baltschekfisky, H. (1992) Inorganic pyrophosphate and inorganic pyrophosphatases. *Molecular Mechanisms in Bioenergetics 23*, 2331–2348

3. Kornberg, A. (1995) Inorganic polyphosphate: toward making a forgotten polymer unforgettable. *J. Bacteriol.* 177, 491–496 CrossRef Medline

4. Kulava, I. S., and Yagabov, V. M. (1983) Polysphosphate metabolism in micro-organisms. *Adv. Microb. Physiol.* 24, 83–171 CrossRef Medline

5. Reeves, R. (1976) How useful is energy in inorganic pyrophosphate. *Trends Biochem. Sci.* 1, 53–55 CrossRef

6. Mertens, E. (1991) Pyrophosphate-dependent phosphofructokinase, an anaerobic glycolytic enzyme? *FEBS Lett.* 285, 1–5 CrossRef Medline

7. Dang, T., and Ingram-Smith, C. (2017) Investigation of pyrophosphate versus ATP substrate selection in the *Entamoeba histolytica* acetate kinase. *Sci. Rep.* 7, 5912 CrossRef Medline

8. Thauer, R. K., Jungermann, K., and Decker, K. (1977) Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.* 41, 100–180 Medline

9. Coombs, G. H., and Müller, M. (1995) Energy metabolism in anaerobic protozoa. *Biochemistry and Molecular Biology of Parasites* 33–47

10. Nagata, R., Fujihashi, M., Sato, T., Atomi, H., and Miki, K. (2018) Identification of a pyrophosphate-dependent kinase and its donor selectivity determinants. *Nat. Commun.* 9, 1765 CrossRef Medline

11. Liu, C.-L., Hart, N., and Peck, H. D., Jr. (1982) Inorganic pyrophosphate: energy source for sulfate-reducing bacteria of the genus *Desulfotomaculum*. *Science* 217, 363–364 CrossRef Medline

12. Carlisle, S. M., Blakeley, S. D., Hemmingsen, S. M., Trevanian, S. J., Hiyoshi, T., Kruger, N. J., and Dennis, D. T. (1990) Pyrophosphate-dependent phosphofructokinase. Conservation of protein sequence between the \( \alpha \) and \( \beta \)-subunits and with the ATP-dependent phosphofructokinase. *J. Biol. Chem.* 265, 18607–1871 Medline

13. Baptiste, E., Moreira, D., and Philippe, H. (2003) Rampant horizontal gene transfer and phospho-donor change in the evolution of the phosphofructokinase. *Gene* 318, 185–191 CrossRef Medline

14. Chi, A., and Kemp, R. G. (2000) The primordial high energy compound: ATP or inorganic pyrophosphate? *J. Biol. Chem.* 275, 35677–35679 CrossRef Medline

15. Alves, A. M., Meijer, W. G., Vrijbloed, J. W., and Dijkhuizen, L. (1996) Characterization and phylogeny of the pfp gene of *Anycylophiopsis methanolica* encoding PPI-dependent phosphofructokinase. *J. Bacteriol.* 178, 149–155 CrossRef Medline

16. Müller, M., Lee, J. A., Gordon, P., Gaasterland, T., and Sensen, C. W. (2001) Presence of prokaryotic and eukaryotic species in all subgroups of the PPI-dependent group II phosphofructokinase protein family. *J. Bacteriol.* 183, 6714–6716 CrossRef Medline

17. Reeves, R. E., and Guthrie, J. D. (1975) Acetate kinase (pyrophosphate): a fourth pyrophosphate-dependent kinase from *Entamoeba histolytica*. *Biochem. Biophys. Res. Commun.* 66, 1389–1395 CrossRef Medline

18. Ingram-Smith, C., Martin, S. R., and Smith, K. S. (2006) Acetate kinase: not just a bacterial enzyme. *Trends Microbiol.* 14, 249–253 CrossRef Medline

19. Fowler, M. L., Ingram-Smith, C., and Smith, K. S. (2012) Novel pyrophosphate-forming acetate kinase from the protist *Entamoeba histolytica*. *Eukaryot. Cell* 11, 1249–1256 CrossRef Medline

20. Matte, A., Tari, L. W., Goldie, H., and Delbaere, L. T. (1997) Structure and mechanism of phosphoenolpyruvate carboxykinase. *J. Biol. Chem.* 272, 8105–8108 CrossRef Medline

21. Matte, A., Goldie, H., Sweet, R. M., and Delbaere, L. T. (1996) Crystal structure of *Escherichia coli* phosphoenolpyruvate carboxykinase: a new structural family with the P-loop nucleoside triphosphate hydrolase fold. *J. Mol. Biol.* 256, 126–143 CrossRef Medline

22. Dunton, P., Belunis, C., Crowther, R., Holfelder, K., Kammlott, U., Levin, W., Michel, H., Ramsey, G. B., Swain, A., Weber, D., and Wertheimer, S. J. (2002) Crystal structure of human cytosolic phosphoenolpyruvate carboxykinase reveals a new GTP-binding site. *J. Mol. Biol.* 316, 257–264 CrossRef Medline

23. Siu, P. M., Wood, H. G., and Stjerneholm, R. L. (1961) Fixation of CO\(_2\) by phosphoenolpyruvic carboxylasephosphorylase. *J. Biol. Chem.* 236, PC21–PC22

24. Siu, P. M., and Wood, H. G. (1962) Phosphoenolpyruvic carboxylasephosphorylase, a CO\(_2\) fixation enzyme from propionic acid bacteria. *J. Biol. Chem.* 237, 3044–3051 Medline

25. Lochmüller, H., Wood, H. G., and Davis, J. J. (1966) Phosphoenolpyruvic carboxylasephosphorylase II: crystallization and properties. *J. Biol. Chem.* 241, 5678–5691 Medline

26. Reeves, R. E. (1970) Phosphoenolpyruvate carboxylase from *Entamoeba histolytica*. *Biochim. Biophys. Acta* 220, 346–349 CrossRef Medline

27. Haberland, M. E., Willard, J. M., and Wood, H. G. (1972) Phosphoenolpyruvate carboxylasephosphorylase: VI: catalytic and physical structures. *Biochemistry* 11, 712–722 CrossRef Medline

28. Chiba, Y., Kamikawa, R., Nakada-Tsukui, K., Saito-Nakano, Y., and Nozaki, T. (2015) Discovery of PPI-type phosphoenolpyruvate carboxykinase
Structural comparisons of phosphoenolpyruvate carboxykinases

genesis in eukaryotes and bacteria. J. Biol. Chem. 290, 23960–23970 CrossRef Medline

29. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797 CrossRef Medline

30. Holm, L., and Laakso, L. M. (2016) Dali server update. Nucleic Acids Res. 44, W351–W355 CrossRef Medline

31. Tari, L. W., Matte, A., Goldie, H., and Delbaere, L. T. (1997) Mg2+ – Mn2+ clusters in enzyme-catalyzed phosphoryl-transfer reactions. Nat. Struct. Mol. Biol. 4, 990–994 CrossRef Medline

32. Mahajan, A., Yuan, C., Lee, H., Chen, E. S., Wu, P.-Y., and Tsai, M.-D. (2008) Structure and function of the phosphothereonine-specific FHA domain. Sci. Signal. 1, re12–re12 Medline

33. Kosinski, J., Mosalaganti, S., von Appen, A., Teimer, R., DiGuilio, A. L., Wang, W., Bui, K. H., Hagen, W. J., Briggs, J. A., Glavy, J. S., Hurt, E., and Beck, M. (2016) Molecular architecture of the inner ring scaffold of the human nuclear pore complex. Science 352, 363–365 CrossRef Medline

34. Park, Y. B., Hohl, M., Padjasek, M., Jeong, E., Jin, K. S., Krel, A., Petrin, J. H., and Cho, Y. (2017) Eukaryotic Rad50 functions as a rod-shaped dimer. Nat. Struct. Mol. Biol. 24, 248–257 CrossRef Medline

35. Iwabuchi, N., Maejima, K., Kitazawa, Y., Ishikawa, H., Nishikawa, M., Tokuda, R., Koinuma, H., Miyazaki, A., Nijo, T., Oshima, K., Yamaji, Y., and Namba, S. (2019) Crystal structure of phyllophy, a phylloidy-inducing effector protein of phytplasmas. Biochem. Biophys. Res. Commun. 513, 952–957 CrossRef Medline

36. Yue, P., Zhang, Y., Mei, K., Wang, S., Lesigang, J., Zhu, Y., Dong, G., and Guo, W. (2017) Sec3 promotes the initial binary t-SNARE complex assembly and membrane fusion. Nat. Commun. 8, 14236 CrossRef Medline

37. Biou, V., Yaremchuk, A., Tukalo, M., and Cusack, S. (1994) The 2.9 Å crystal structure of T. thermophilus seryl-tRNA synthetase complexed with tRNA<sup>ser</sup>. Science 263, 1404–1410 CrossRef Medline

38. Tari, L. W., Matte, A., Pugazhenthii, U., Goldie, H., and Delbaere, L. T. (1996) Snapshot of an enzyme reaction intermediate in the structure of the ATP–Mg<sup>2+</sup>– oxalate ternary complex of Escherichia coli PEP carboxykinase. Nat. Struct. Mol. Biol. 3, 355–363 CrossRef Medline

39. Sullivan, S. M., and Holyoak, T. (2007) Structures of rat cytosolic PEPCK: insight into the mechanism of phosphorylation and decarboxylation of oxaloacetic acid. Biochemistry 46, 10078–10088 CrossRef Medline

40. Carlson, G. M., and Holyoak, T. (2009) Structural insights into the mechanism of phosphoenolpyruvate carboxykinase catalysis. J. Biol. Chem. 284, 27037–27041 CrossRef Medline

41. Delbaere, L. T., Sudom, A. M., Prasad, L., Leduc, Y., and Goldie, H. (2004) Structure/function studies of phosphoryl transfer by phosphoenolpyruvate carboxykinase. Biochim. Biophys. Acta 1697, 271–278 CrossRef Medline

42. Fukuda, W., Fukui, T., Atomi, H., and Imanaka, T. (2004) First characterization of an archaeal GTP-dependent phosphoenolpyruvate carboxykinase from the hyperthermophilic archaean Thermococcus kodakaraensis KOD1. J. Bacteriol. 186, 4620–4627 CrossRef Medline

43. Holyoak, T., Sullivan, S. M., and Nowak, T. (2006) Structural insights into the mechanism of PEPCK catalysis. Biochemistry 45, 8254–8263 CrossRef Medline

44. Cotelesage, J. J., Prasad, L., Zeikus, J. G., Laivenieks, M., and Delbaere, L. T. (2005) Crystal structure of Anaerobiospirillum succiniciproducens PEP carboxykinase reveals an important active site loop. Int. J. Biochem. Cell Biol. 37, 1829–1837 CrossRef Medline

45. Trapani, S., Linss, J., Goldenberg, S., Fischer, H., Craievich, A. F., and Oliva, G. (2001) Crystal structure of the dimeric phosphoenolpyruvate carboxykinase (PEPCK) from Trypanosoma cruzi at 2 Å resolution. J. Mol. Biol. 313, 1059–1072 CrossRef Medline

46. Jurado, L. A., Machín, I., and Urbina, J. A. (1996) Trypanosoma cruzi phospho enol pyruvate carboxykinase (ATP-dependent): transition metal ion requirement for activity and sulfhydryl group reactivity. Biochim. Biophys. Acta 1292, 188–196 CrossRef Medline

47. Nakatsu, T., Kato, H., and Oda, J. (1998) Crystal structure of asparagine synthetase reveals a close evolutionary relationship to class II aminocyl-tRNA synthetase. Nat. Struct. Biol. 5, 15–19 CrossRef Medline

48. Perutz, M. F., Kendrew, J. C., and Watson, H. C. (1965) Structure and function of haemoglobin: II: some relations between polypeptide chain configuration and amino acid sequence. J. Mol. Biol. 13, 669–678 CrossRef

49. Babbitt, P. C., and Gerlt, J. A. (1997) Understanding enzyme superfamilies chemistry as the fundamental determinant in the evolution of new catalytic activities. J. Biol. Chem. 272, 30591–30594 CrossRef Medline

50. Hanefeld, U., Gardossi, L., and Magner, E. (2009) Understanding enzyme immobilisation. Chem. Soc. Rev. 38, 453–468 CrossRef Medline

51. Kabach, W. (2010) XDS: Xds. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132 CrossRef Medline

52. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. 69, 1204–1214 CrossRef Medline

53. Terviliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H., and Hung, L.-W. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. Acta Crystallogr. D Biol. Crystallogr. 65, 582–601 CrossRef Medline

54. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 CrossRef Medline

55. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Loerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. I., Sacchettini, J. C., Sauter, N. K., and Terviliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. 58, 1948–1954 CrossRef Medline

56. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. 67, 355–367 CrossRef Medline

57. Sullivan, S. M., and Holyoak, T. (2008) Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection. Proc. Natl. Acad. Sci. U.S.A. 105, 13829–13834 CrossRef Medline

58. Waterhouse, A., Bertoni, M., Biennert, S., Studer, G., Tauriello, G., Gumienny, R., Feer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296–W303 CrossRef Medline

59. Roberts, E., Eargle, J., Wright, D., and Luthney-Shulten, Z. (2006) MultiSeq: unifying sequence and structure data for evolutionary analysis. BMC Bioinformatics 7, 382 CrossRef Medline

60. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. J. Mol. Graphics 14, 33–38, 27, 28 CrossRef Medline

61. Roberts, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324 CrossRef Medline