Lysine Biosynthesis of *Thermococcus kodakarensis* with the Capacity to Function as an Ornithine Biosynthetic System*

Received for publication, June 12, 2016, and in revised form, August 24, 2016 Published, JBC Papers in Press, August 26, 2016, DOI 10.1074/jbc.M116.743021

Ayako Yoshida‡, Takeo Tomita‡, Haruyuki Atomi§, Tomohisa Kuzuyama‡, and Makoto Nishiyama‡

From the ‡Biotechnology Research Center, University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo 113-8657, the §Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Katsura, Nishikyo-ku, Kyoto 615-8510, and the ‡Japan Science and Technology Agency, CREST, 7, Gobancho, Chiyoda-ku, Tokyo 102-0076 Japan

We recently discovered a biosynthetic system using a novel amino group carrier protein called LysW for lysine biosynthesis via α-aminoadipate (AAA), and revealed that this system is also utilized in the biosynthesis of arginine by *Sulfolobus*. In the present study, we focused on the biosynthesis of lysine and ornithine in the hyperthermophilic archaeon *Thermococcus kodakarensis*, and showed that their biosynthesis is accomplished by a single set of metabolic enzymes. We also determined the crystal structure of the LysX family protein from *T. kodakarensis*, which catalyzes the conjugation of LysW with either AAA or glutamate, in a complex with LysW-γ-AAA. This crystal structure is the first example to show how LysX recognizes AAA as a substrate and provides a structural basis for the bifunctionality of the LysX family protein from *T. kodakarensis*. Based on comparisons with other LysX family proteins, we propose a mechanism for substrate recognition and its relationship with molecular evolution among LysX family proteins, which have different substrate specificities.

Metabolic pathways are considered to have evolved in response to the depletion of compounds used for the growth of heterotrophic organisms. In ancient organisms, metabolic pathways comprised enzymes with broad substrate specificities due to the small sizes of their genomes (1, 2), in which a single enzyme with broad substrate specificity had been recruited for several similar reactions in multiple metabolic processes to utilize the compounds necessary for growth. In the process of evolution, genes encoding ancestral-type enzymes, which have the ability to function in multiple pathways, were duplicated and differentiated to encode the enzymes specialized for each metabolic pathway. Consequently, in extant organisms, there are several homologous metabolic pathways, in which corresponding reactions are catalyzed by homologous enzymes, as observed in leucine and isoleucine biosynthesis (the pyruvate pathway) as well as in histidine, serine, and tryptophan biosynthesis (1, 3). The enzymes in each biosynthetic pathway are considered to share the same enzyme origin and evolved by gene duplication and specialization. Previous studies based on structural biology also supported this hypothesis because a large number of proteins with similar catalytic mechanisms were found to share the same protein-fold, and, thus, may have evolved from a common scaffold (4). We previously discovered a novel type of biosynthetic pathway for lysine using α-aminoadipate (AAA)2 as an intermediate in the thermophilic bacterium *Thermus thermophilus* (5–7). Although lysine is generally biosynthesized from aspartate through diaminopimelate (DAP) in bacteria (8), *T. thermophilus* synthesized lysine through AAA. This lysine biosynthetic AAA pathway is composed of two parts. The former part, the conversion of 2-oxoglutarate to AAA, is the same as the corresponding part in fungal lysine biosynthesis via AAA, and consists of homocitrate synthase, (homo)aconitase, homoisocitrate dehydrogenase, and AAA aminotransferase. These enzymes are homologous to enzymes involved in leucine biosynthesis and a part of the tricarboxylic acid cycle (7, 9, 10). The latter part of this novel lysine biosynthetic pathway, the conversion from AAA to lysine, is accomplished by five enzymes: LysX, LysZ, LysY, LysJ, and LysK, using the amino group carrier protein LysW. The α-amino group of AAA is modified by LysW by LysX, forming an isopeptide bond between the α-amino group of AAA and γ-carboxyl group of the glutamate residue at the C terminus of LysW. LysW-γ-AAA thus synthesized is then transferred to subsequent biosynthetic enzymes to be converted to LysW-γ-lysine by phosphorylation, reduction, and amination steps. In the final step, LysW-γ-lysine is recognized by LysK, a carboxypeptidase, resulting in the release of lysine. LysW contains many acidic amino acid residues for electrostatic interactions with each enzyme, and, thus, functions as an amino group carrier protein for efficient lysine biosynthesis (11, 12). Since the reactions in this lysine AAA pathway are similar to those in the glutamate to ornithine conversion in arginine biosynthesis, the latter parts of the lysine AAA pathway and arginine biosynthetic pathway may have a common evolutionary origin; however, the first reaction step catalyzed by LysX differs from that in arginine biosynthesis in which the α-amino group of glutamate is modified by N-acetylation (13–15). We previously revealed that the hyperthermopoacidophilic archaeon *Sulfolobus acidocaldarius*, biosynthesizes lysine and arginine via a LysW-

---

*This work was supported in part by Japan Society for the Promotion of Science KAKENHI Grant 24228001 (to M. N.). The authors declare that they have no conflicts of interest with the contents of this article. The atomic coordinates and structure factors (code 5K2M) have been deposited in the Protein Data Bank ([http://wwpdb.org/](http://wwpdb.org/)).

‡1 To whom correspondence should be addressed. Tel.: 81-3-5841-3074; Fax: 81-3-5841-8030; E-mail: umanis@mail.ecc.u-tokyo.ac.jp.

‡2 The abbreviations used are: AAA, α-aminoadipate; TkLysX/ArgX, bifunctional LysX/ArgX from *T. kodakarensis*; TkLysW, LysW from *T. kodakarensis*; TrLysX, LysX from *T. thermophilus*; TrLysW, LysW from *T. thermophilus*; StArgX, ArgX from *S. tokodaii*; StLysW, LysW from *S. tokodaii*; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; PDB, Protein Data Bank; DAP, diaminopimelate; AMP-PNP, 5′-adenylyl-β,γ-imidodiphosphate.
mediated pathway (11). \textit{S. acidocaldarius} has two genes that code for the LysX family proteins, SaLysX and SaArgX, which catalyze the formation of LysW-\gamma-AAA and LysW-\gamma-glutamate (LysW-\gamma-Glu) used for lysine and arginine biosynthesis, respectively. \textit{S. acidocaldarius} has a single set of biosynthetic enzymes for the subsequent steps (LysW-\gamma-AAS to lysine and LysW-\gamma-Glu to ornithine) on its genome. In \textit{Sulfolobus}, the enzymes LysZ/ArgB, LysY/ArgC, LysJ/ArgD, and LysK/ArgE, which are responsible for conversion processes, are used as bifunctional enzymes for lysine and arginine biosynthesis. This has become an evolutionary trait demonstrating that lysine and arginine biosynthesis enzymes have evolved from common ancestral enzymes.

In the current genome database, we found that some bacteria and archaea utilize the LysW system for lysine and/or arginine biosynthesis. In the present study, we analyzed the LysW-mediated biosynthetic pathway in the hyperthermophilic archaeon \textit{Thermococcus kodakarensis}, which only has a single genomic copy of lysine biosynthetic genes including \textit{lysW}, \textit{lysX}, \textit{lysY}, \textit{lysZ}, \textit{lysJ}, and \textit{lysK}. We found that \textit{T. kodakarensis} has the ability to biosynthesize lysine and ornithine using a completely single set of enzymes including LysX/ArgX. We also examined the crystal structure of the LysX/ArgX in a complex with LysW-\gamma-AAA, which provided the structural basis for LysX/ArgX to exhibit dual functions.

Results

The Lysine Biosynthetic Gene Cluster in \textit{T. kodakarensis}—In the current genome database, the genes encoding the amino group carrier protein, LysW are found in many organisms, which are mainly categorized as thermophilic bacteria and archaea. The hyperthermophilic archaeon \textit{T. kodakarensis} does not have genes for lysine biosynthesis through DAP, but has TK0279 (\textit{lysW} homolog) in the gene cluster for lysine biosynthetic genes through AAA (Fig. 1A), suggesting that it produces lysine with a LysW-mediated system. In this cluster, TK0279, TK0278, TK0276, TK0277, TK0275, and TK0274 were expected to be responsible for the conversion process from AAA to lysine through the LysW system as LysW-\gamma-AAA synthetase, LysW-\gamma-AAA kinase, LysW-\gamma-AAA semialdehyde dehydrogenase, LysW-\gamma-lysine aminotransferase, and LysW-\gamma-lysine carboxypeptidase, respectively (Fig. 1B). On the other hand, \textit{T. kodakarensis} had no other homologous genes for arginine biosynthetic genes.

Characterization of TK0278, a \textit{lysX} Homolog—We investigated whether the lysine biosynthetic enzymes of \textit{T. kodakarensis} converted AAA/Glu to lysine and ornithine. When we characterized the TK0278 (\textit{lysX} homolog) gene product, purified recombinant TK0278 recognized AAA and glutamate as substrates and ligated both compounds with TK0279, an amino group carrier protein LysW in this microorganism, in an ATP-dependent manner. The specific activities of TK0278 for AAA and glutamate were 0.021 ± 0.0005 and 0.016 ± 0.0002 units mg\(^{-1}\), respectively. We hereafter referred to TK0278 as TkllysX/ArgX.

Characterization of TK0276, a \textit{lysZ} Homolog—We characterized TK0276, which was expected to encode an enzyme that catalyzes the second step: the phosphorylation of LysW-\gamma-AAA and LysW-\gamma-Glu formed by TkllysX/ArgX. We previously demonstrated that the LysW-\gamma-AAA phosphate formed by LysZ from \textit{T. thermophilus} exhibited no detectable change in migration from that of LysW-\gamma-AAA, on Tricine SDS-PAGE, whereas it showed slower mobility when it was chemically converted to a stable hydroxamate derivative by a hydroxylamine treatment (12). Therefore, we treated TkllysW (TK0279)-\gamma-AAA or TkllysW-\gamma-Glu, which were prepared by the TkllysX/ArgX reaction, with TK0276 in the presence of ATP, MgSO\(_4\), and hydroxylamine, and applied the reaction mixture to Tricine SDS-PAGE. As a result, a band shift to the upper position was observed for both reaction mixtures using TkllysW-\gamma-AAA and TkllysW-\gamma-Glu as substrates (Fig. 2A), suggesting that TK0276 also functions as a bifunctional enzyme capable of recognizing both substrates.

TkllysW has no lysine or arginine residues. Therefore, the E42R mutation was introduced to TkllysW for the easy detection of the C-terminal adduct after the trypsin treatment. TkllysW(E42R)-\gamma-AAA or TkllysW(E42R)-\gamma-Glu was prepared by TkllysX/ArgX, and each derivative was used for the TK0276 reaction, followed by the chemical conversion of the phosphate group to hydroxamate by hydroxylamine. Because a band shift was also observed for the TkllysW(E42R) mutant derivatives on Tricine SDS-PAGE (Fig. 2B), we extracted these bands from the gel and digested them with trypsin. The C-terminal tryptic fragments (Arg\(^{42}\)-Glu\(^{53}\)) of the TkllysW(E42R) derivatives were analyzed by MALDI-TOF-MS and MS/MS analyses. When intact TkllysW(E42R)-\gamma-AAA was analyzed by MALDI-TOF-MS, a peptide with a molecular mass of 1474.66 Da, which corresponds to the molecular mass of the C-terminal peptide \textit{ELEPEVEEDWGF}\(^{53}\) + AAA (theoretical mass of 1474.61 Da), was obtained (Fig. 3A). On the other hand, when the TkllysW(E42R)-\gamma-AAA derivative that exhibited retarded mobility on Tricine/SDS-PAGE was analyzed, a fragment with a molecular mass 1489.65 Da, corresponding to the mass of the peptide42ELPEVEEDWGE53
peptide with the sequence 42ELPEVEEDWGE53 + AAA-CONHOH (theoretical mass of 1489.62 Da), was found (Fig. 3B). Similarly, an analysis of the intact C-terminal fragment of TkLysW(E42R)-Glu and that was treated with TK0276 followed by chemical conversion gave a molecular mass of 1460.60 Da corresponding to 42ELPEVEEDWGE53 + Glu (Fig. 3C), and a molecular mass of 1475.61 Da corresponding to 42ELPEVEEDWGE53 + Glu-CONHOH (Fig. 3D). In addition, the MS/MS analysis verified that the modification by TK0278 and successive phosphorylation by TK0276 occurred at the C terminus of TkLysW (Fig. 4, A and B). These results indicated that TK0276 also recognizes TkLysW-γ-AAA and TkLysW-γ-Glu as substrates, thereby acting as a bifunctional enzyme. TK0276 was hereafter referred to as TkLysZ/ArgB.

In Vitro Conversion of AAA to Lysine and Glutamate to Ornithine—To clarify whether the latter half of lysine biosynthetic enzymes catalyze the reaction required for ornithine biosynthesis, we reconstituted the reactions involved in the biosynthesis of lysine and ornithine from AAA and glutamate, respectively, in vitro. We purified TkLysW, TkLysX/ArgX (TK0278), TkLysZ/ArgB (TK0276), TK0277, TK0275, and TK0274 from Escherichia coli cells (Fig. 5A). All proteins were mixed with their putative cofactors in the presence of AAA or glutamate and then incubated at 60 °C for 1 h. When AAA was used as a substrate, lysine synthesis (93 μM) was observed (Fig. 5B). Moreover, we observed the production of ornithine (105 μM) when using glutamate as the starter compound (Fig. 5C). When the reaction mixtures lacking any proteins were analyzed, the production of neither lysine nor ornithine was observed. These results indicated that all six proteins were necessary and sufficient for the biosynthesis of lysine and ornithine, and TK0277, TK0275, and TK0274 functioned as bifunctional TkLysY/ArgC, TkLysJ/ArgD, and TkLysK/ArgE, respectively. Collectively, these results demonstrated that T. kodakarensis has the potential to biosynthesize lysine and ornithine using the LysW-mediated system with a single set of bifunctional enzymes.

Crystal Structure of TkLysX/ArgX Complexed with TkLysW-γ-AAA—To elucidate the substrate-recognition mechanism of bifunctional TkLysX/ArgX, we crystallized TkLysX/ArgX in...
the presence of TkLysW and AMP-PNP, MgSO₄, and AAA or glutamate. As a result, we obtained the crystal structure of TkLysX/ArgX complexed with TkLysW-H₉₂₅₃-AAA (TkLysX/ArgX-H₁₈₅₂₈₈₇-LysW-H₉₂₅₃-AAA) at a resolution of 2.18 Å, which corresponded to the structure of the post-reaction state of TkLysX/ArgX. However, we did not obtain good crystals containing TkLysW-H₉₂₅₃-Glu. Data collection and refinement statistics are summarized in Table 1. The asymmetric unit contained two TkLysX/ArgX tetramers, two intact TkLysW monomers each with one zinc atom, four partial TkLysW, eight ADP molecules, five phosphate ions, two sulfate ions, nine magnesium atoms, one free AAA molecule, and 628 water molecules. Although we

FIGURE 4. LC-MS/MS analysis of trypsin-digested TkLysW derivatives. A, TkLysW-AAA-NHOH prepared by the incubation of TkLysW-γ-AAA with TK0276, ATP, and hydroxylamine. B, TkLysW-Glu-NHOH prepared by the incubation of TkLysW-γ-Glu with Tk0276, ATP, and hydroxylamine. Theoretical m/z values of the fragments were provided.
added AMP-PNP to the crystallization drops, there was no electron density for the γ-phosphate group of AMP-PNP. Therefore, we assigned ADP for density. One TkLysX/ArgX tetramer (chains A-D, tetramer 1) formed a complex with two TkLysW chains (chains E and F), whereas another TkLysX/ArgX tetramer (chains G-J, tetramer 2) in the asymmetric unit may bind four TkLysW (chains K-N), the electron densities of which were only partially observed. In tetramer 2, we observed the partial densities of globular domains and C-terminal extensions for chains L and N, and only the C-terminal extensions for chains K and M (Fig. 6, A–C). In the crystal structure of the ArgX and LysW complex from Sulfolobus tokodaii (StArgX/HL18528 LysW), the StArgX tetramer interacts with two StLysW, indicating that two of four active sites are vacant (11); therefore, we speculated that tetramer 1 is the functional unit and tetramer 2 is the crystallographic artifact. We hereafter only discussed the structure of tetramer 1, which was composed of the TkLysX/ArgX tetramer binding two TkLysW.

The structure of TkArgX/LysX was similar to those of LysX from T. thermophilus (TtLysX) and StArgX (root mean square deviation values of 1.34–1.51 and 1.46–1.73 Å, respectively), and comprised three domains: domains A (residues 1–93), B (residues 94–167), and C (residues 168–273), as observed in ATP-grasp family enzymes. TkLysW is composed of an N-terminal globular domain (residues 1–43) and C-terminal extension (residues 44–53). The globular domain takes a zinc-finger fold, similar to StLysW and LysW from T. thermophilus (TtLysW) (12), coordinating a zinc atom with four cysteine residues (Cys-4, -7, -24, and -27).

In the crystal structure, the C terminus of TkLysW was inserted into the catalytic site of TkLysX/ArgX. AAA was covalently bound to the C-terminal Glu residue of TkLysW, in which the α-amino group of AAA was connected to the γ-carboxyl group of Glu forming an isopeptide bond (Fig. 6D). Moreover, ADP molecules and phosphate ions (or sulfate ions, which may come from the crystallization solution) were present.
TABLE 1

| Data collection and refinement statistics |
|------------------------------------------|
| TkLysX/ArgX-TkLysW-γ-AAA                 |

| Data collection* | TkLysX/ArgX-TkLysW-γ-AAA |
|------------------|--------------------------|
| X-ray source     | PF-AR NE3A               |
| Wavelength (Å)   | 1.000                    |
| Space group      | P2₁                    |
| Cell dimensions (Å) | a = 76.2, b = 135.1, c = 137.0, β = 97.5° |
| Resolution (Å)  | 2.18 (2.22–2.18)        |
| Reflections (total/unique) | 534,873/142,143 |
| Riref (%)        | 8.7 (38.8)              |
| Rmerge (%)       | 4.5 (20.2)              |
| CC1/2            | 99.6 (87.5)             |
| I/σ (f)          | 18.6 (2.59)             |
| Completeness (%) | 99.7 (100.0)            |

* Values in parentheses are data for the highest resolution shell.

R-factor (work/test) (Å) | 18.2/22.0 |
No. of atoms | 19,516 |
Protein atoms | 18,560 |
ADP | 216 |
SO₄²⁻ | 10 |
PO₄³⁻ | 25 |
Mg²⁺ | 9 |
Zn²⁺ | 2 |
AAA | 66 |
Water | 628 |
Average B-factor (Å²) | 43.4 |
Protein atoms | 43.1 |
ADP | 38.9 |
SO₄²⁻ | 85.3 |
PO₄³⁻ | 49.4 |
Mg²⁺ | 47.8 |
Zn²⁺ | 55.8 |
AAA | 47.6 |
Water | 51.4 |
Root mean square deviation values |
| Bond length (Å) | 0.008 |
| Bond angle (°) | 1.4 |
| Ramachandran plot² |
| Favored (%) | 97.9 |
| Allowed (%) | 2.1 |
| Outlier (%) | 0.1 |

² Data were calculated using MolProbity (30).

“GSWGR” motif (residues 126–130) in the β7–β8 loop, which is also conserved and involved in the recognition of the C-terminal extension of LysW in StArgX, mostly through hydrogen bonding (Fig. 7C). These results clearly demonstrate that negatively charged TkLysW is attracted by TkLysX/ArgX electrostatically; however, the interaction to orient the C-terminal extension of TkLysW correctly at the catalytic site of TkLysX/ArgX is stabilized by a hydrogen bond interaction between two proteins. We speculate that the weak interaction via hydrogen bonds facilitates the translocation of LysW derivatives to an enzyme catalyzing the next reaction.

Active Site Structure and Reaction Mechanism—ADP, a phosphate ion, two magnesium atoms, and the C terminus of TkLysW-γ-AAA were located in the active site of TkLysX/ArgX (Figs. 6D and 8A). The α-carboxyl group of AAA conjugated to the C terminus of TkLysW was recognized by Thr¹⁹⁶-N, Thr¹⁹⁶-O⁶, Ala¹⁹⁷-N, and Arg¹⁸⁷-N⁷¹, whereas the δ-carboxyl group interacted with Thr¹⁹⁶-O⁴, Asn²⁵⁰-N, Asn²⁵⁰-N⁶², and Ala²⁵¹-N (Fig. 8B). Loop β₁₁–β₁₂ (large loop: residues 189–200), containing Thr¹⁹⁶ and Ala¹⁹⁷, covered the catalytic site together with loop β₇–β₈ (small loop: residues 123–132) (Fig. 8A). Several residues from large and small loops were responsible for the interaction with the ribose and phosphate moieties of ADP, respectively (Fig. 8C). The inorganic phosphate ion (or sulfate ion) was bound between the β-phosphate group of ADP and α-amino group of AAA connected with the γ-carboxyl group of Glu⁵³ from TkLysW. In the ATP-dependent carboxylate-amine ligase family, the reaction is initiated from the activation of the carboxyl group by phosphorylation, and the amino group of the other substrate then nucleophilically attacks the phosphorylated carbon atom to form the peptide bond (16). When the catalytic site was compared between the TkLysX/ArgX-TkLysW-γ-AAA:ADP and StArgX:StLysW:Glu-ADP complexes, the residues involved in AAA/Glu and ADP binding were conserved and superposed well with each other (Fig. 8D), suggesting that the catalytic mechanism of TkLysX/ArgX is the same as that of StArgX. On the other hand, the residues (Thr¹⁹⁶, Asn²⁵⁰, and Ala²⁵¹) involved in the recognition of the δ-carboxyl group of AAA were not conserved in StArgX, which recognizes glutamate, but not AAA as a substrate as described below, implying that these residues are responsible for differences in substrate specificity between monofunctional StArgX and bifunctional TkLysX/ArgX.

Substrate-recognition Mechanism of TK0278—LysX family proteins may be divided into three groups based on their substrate specificities: 1) AAA-specific LysX responsible for lysine biosynthesis; 2) glutamate-specific ArgX for arginine biosynthesis; and 3) bifunctional LysX/ArgX for lysine and arginine biosynthesis such as TkLysX/ArgX. In our previous study, we examined the crystal structure of StArgX and elucidated the mechanism of glutamate recognition (11). Based on the recognition mechanism in StArgX and a comparison of amino acid sequences among LysX family proteins, we proposed the signature motif located at the N terminus of helix α₈ (corresponding to helix α⁹ in TkLysX/ArgX) (the Asn-Ser/Thr/Val motif in LysX and Gly-Leu/Phe and Ala-Leu motifs in ArgX) that defined substrate specificity (Fig. 9A). In StArgX, the main chain nitrogen atoms of the Gly-Phe motif at the N terminus of

in the catalytic site, suggesting that this crystal structure is in the post-reaction state.

Interaction between TkLysX/ArgX and TkLysW-γ-AAA—We previously revealed that LysW functions not only as the N-modifying group, but also as the carrier protein interacting with biosynthetic enzymes (11, 12). When the electrostatic potential was mapped to the surface of the TkLysX/ArgX-TkLysW-γ-AAA structure, TkLysW-γ-AAA was covered by a negatively charged surface, and the area around the active site of TkLysX/ArgX was positively charged, suggesting that TkLysW interacts with TkLysX/ArgX electrostatically (Fig. 7A). Thus, it is plausible that LysW functions as an amino group carrier protein in the lysine and arginine biosynthetic pathways in T. kodakarenensis as well as in those in T. thermophilus and S. tokodaii. However, the N-terminal globular domain of TkLysW was recognized by TkLysX/ArgX through a limited number of ionic and hydrogen bond interactions, as was the case for the StArgX:LysW-Glu-ADP complex (11) (Fig. 7B).

On the other hand, residues in the C-terminal extension of TkLysW, particularly the conserved C-terminal “EDWGE” motif (residues 49–53), were strictly recognized by the...
helix α8 recognized the γ-carboxyl moiety of glutamate (Fig. 9B), and the corresponding residues in LysX were conserved as Asn-Thr/Ser to bind the δ-carboxyl group of AAA (Fig. 9, A and C). In the TkLysX/ArgX-TkLysW-γ-AAA structure, the δ-carboxyl group of AAA interacted with Asn and Ala in the signature motif that we proposed. In addition to these residues, the δ-carboxyl group of AAA was recognized by the hydrogen bond between Thr-O\(^\dagger\) (Fig. 9D). Amino acid sequence alignment among LysX family proteins showed that this threonine residue was conserved in LysX and bifunctional LysX/ArgX, but not in ArgX (Fig. 9A), suggesting that this threonine is important for the recognition of AAA. In our previous study, Gly-Phe, which is replaced with Asn-Ala in TkLysX/ArgX, was proposed to be the determinant of substrate specificity of StArgX based on the finding that replacement of the Asn-Thr sequence to Gly-Phe converted LysX from *S. acidocaldarius* to an enzyme that recognized glutamate as a substrate and exhibited decreased activity as LysX; however, the γ-carboxyl group of glutamate was also recognized by Tyr in StArgX (Fig. 9B). Notably, Tyr was conserved well in ArgX, whereas it was replaced with isoleucine in LysX and bifunctional LysX/ArgX (Fig. 9A), suggesting that Tyr is also...
present in bifunctional LysX/ArgX, we expect Tyr175 to be a member of the residues comprising the signature motif to recognize glutamate. In TkLysX/ArgX lacking the tyrosine residue at the corresponding position (185), another tyrosine residue (Tyr^{175}) extended its side chain toward the substrate-binding site. Although this Tyr^{175} did not interact with AAA, it was in a position that allowed an interaction with the γ-carboxyl group of bound glutamate with a shorter methylene moiety than that of AAA (Fig. 9). The interaction between the globular domain of TkLysW (chain E, salmon) and TkLysX/ArgX (chain A (green) and chain D (yellow)). The residues involved in the interaction are shown in the stick model. C, recognition of the C-terminal extension of TkLysW-γ-AAA (chain E, salmon) by TkLysW/ArgX (chain A (green) and chain D (magenta)).

**Discussion**

The patchwork hypothesis of evolution in metabolic biosynthetic pathways suggests that gene duplication and functional differentiation occurred for genes encoding enzymes with multifunctions, thereby allowing these enzymes/pathways to evolve into more specific ones. In the present study, we exemplified the existence of multifunctional amino acid biosynthetic pathways in the hyperthermophilic archaeon *T. kodakarensis*. We verified in vitro that lysine and ornithine were both produced from AAA and glutamate, respectively, by a single set of enzymes (TK0274-TK0279) using a LysW-mediated biosynthetic system. Because the gene cluster TK0280-TK0283, which is putatively involved in the first half of lysine biosynthesis, is present upstream of TK0274-TK0278 in *T. kodakarensis*, it is plausible to expect *T. kodakarensis* to biosynthesize lysine from 2-oxoglutarate through AAA, similar to *T. thermodilophus*. Moreover, *T. kodakarensis* has no other homologous genes responsible for the conversion of oxaloacetate to 2-oxoglutarate, which may be converted to glutamate needed for transamination to synthesize amino acids. Therefore, we consider TK0280-TK0283 to be involved in the oxaloacetate to 2-oxoglutarate conversion as well as lysine biosynthesis. Taken together, this genome information suggests that ornithine is also produced by the same set of enzymes from oxaloacetate in a similar manner to that for lysine biosynthesis in vivo. We previously demonstrated that *Sulfolobus* biosynthesized lysine and arginine using the LysW system (11), in which *Sulfolobus* has two LysX family proteins, LysX and ArgX, specifically catalyzing the reactions for lysine and arginine biosynthesis. Another four enzymes were used for lysine and ornithine biosynthesis as bifunctional enzymes. *T. kodakarensis* only has a single LysX family protein, TK0278, which was shown to be capable of acting as a bifunctional enzyme in this study. We also characterized TK0277, corresponding to LysZ in lysine biosynthesis and indicated that TK0277 had the ability to catalyze the phosphorylation of LysW-γ-AAA and LysW-γ-Glu, which are the products of TkLysX/ArgX. These results coupled with the in vitro biosynthesis of lysine and ornithine using purified enzymes/proteins suggest that *T. kodakarensis* has ancient-type lysine and ornithine biosynthetic pathways composed of a set of bifunctional enzymes. The promiscuity and/or multi-

![FIGURE 7. Interaction between TkLysX/ArgX and TkLysW-γ-AAA. A, electrostatic potential of the TkLysW-γ-AAA and TkLysX/ArgX surfaces of the complex. The yellow circle indicates the binding site of TkLysW-γ-AAA on TkLysX/ArgX. B, interaction between the globular domain of TkLysW (chain E, salmon) and TkLysX/ArgX (chain A (green) and chain D (yellow)). The residues involved in the interaction are shown in the stick model. C, recognition of the C-terminal extension of TkLysW-γ-AAA (chain E, salmon) by TkLysW/ArgX (chain A (green) and chain D (magenta)).](Image)
functionality of proteins is a key feature in enzyme evolution because the functions of promiscuous enzymes may be easily altered by even a single mutation to adapt to the new environment (17). In the present study, we found that several mutations in TkLysX/ArgX altered substrate preferences toward AAA or glutamate, indicating that the substrate specificities of bifunctional enzymes may be changed toward a specific function by a limited number of mutations. We speculate that this pathway in

FIGURE 8. The active site structure of TkLysX/ArgX. A, an overview of the active site. Green and pale pink chains indicate chain A of the TkLysX/ArgX tetramer and TkLysW-γ-AAA (chain E), respectively. The C-terminal extension loop of TkLysW (Glu49-Glu53), AAA, ADP (blue), and PO4\(^2^-\) are shown in the stick model. Two Mg\(^{2+}\) ions are depicted in the non-bonded sphere models. Large and small loops are colored in orange. B, AAA-binding site. AAA and Glu53 of TkLysW-γ-AAA (chain E) are shown in pale pink and residues in green are from TkLysX/ArgX. C, stereo view of residues involved in the recognition of ADP, phosphate, and the Mg ion. ADP is shown in the blue stick model. D, superposition of the catalytic site between TkLysX/ArgX and StArgX. Residues from TkLysX/ArgX and StArgX are shown in green and cyan, respectively (stereo view).
T. kodakarensis still retains the evolutionary trace in the lysine/ornithine (arginine) biosynthetic pathways before gene duplication and successive differentiation events.

However, T. kodakarensis, which belongs to Thermococcales in Euryarchaeota, does not grow in the absence of arginine (18), implying that this strain cannot biosynthesize arginine in vivo. Arginine auxotrophy may be due to the lack of argG and argH genes, which encode argininosuccinate synthase and argininosuccinate lyase, respectively. A gene cluster similar to that of T. kodakarensis has been found in other Thermococcus strain. Thermococcus paralvinellae is the only organism that has a gene cluster composed of 14 genes (TES1_0498 to TES1_0511) including the argG, argH, carA, and carB homologs in 13 Thermococcus strains in the KEGG database (19) (Fig. 11). In addition to this gene cluster, T. paralvinellae has the argF homolog encoding ornithine carbamoyltransferase in another locus (TES1_0515), suggesting that T. paralvinellae has the ability to convert ornithine to arginine, although we did not know the amino acid requirements of this strain. Although most Pyrococcus strains, also belonging to Thermococcales, have the same lysine biosynthetic gene cluster as the lysW homolog and lack other genes for arginine biosynthesis, some Pyrococcus species (Pyrococcus furiosus COM1, P. furiosus DSM3638, Pyrococcus sp. ST04, and Pyrococcus sp. NA2) have very interesting gene

FIGURE 9. Substrate recognition mechanism in LysX family proteins. A, amino acid sequence alignment of LysX family proteins. Residues related to substrate specificity are indicated by black rectangles, and the residue numbers of TkLysX/ArgX are written below. TTC1543 (TtLysX), DR_2194, STK_01920, SAC_0754, and SAC_1621, LysX were from T. thermophilus, Tp_2194, PT_01920, SA_0754, and SA_1621, respectively. A gene cluster similar to that of T. kodakarensis, which has the same lysine biosynthetic gene cluster as the lysW homolog and lack other genes for arginine biosynthesis, some Pyrococcus species (Pyrococcus furiosus COM1, P. furiosus DSM3638, Pyrococcus sp. ST04, and Pyrococcus sp. NA2) have very interesting gene
Lysine Biosynthesis of T. kodakarensis

organisms. These strains possess an entire lysine biosynthetic gene cluster including lysW. Moreover, they carry a gene cluster containing argG and argH to biosynthesize arginine from ornithine on a different genome locus. The cluster on the different locus also contains an additional lysX homolog. The lysX homologs in the lysine biosynthetic gene cluster encode enzymes that possess the signature motif (Tyr175, Ile185, Thr196, and Asn250–Ala251) for the LysX/ArgX bifunctional enzyme. The signature motif of bifunctional LysX/ArgX is similar to that of LysX; however, Tyr at position 175 allows bifunctional LysX/ArgX to recognize glutamate and AAA as substrates. On the other hand, LysX family proteins (PFC_00105, PF0209, Py04_0364, and PNA2_0499) do not have the typical signature motif conserved in bifunctional LysX/ArgX, as observed in TkLysX/ArgX, but have Ile, Ile, Ala, and Leu/Ser-Ala (Tyr, Ile, Met, and Asn-Thr in PNA2_0499) at positions corresponding to 175, 185, 196, and 250–251 in TkLysX/ArgX, respectively (Fig. 9A). Three (PFC_00105, PF0209, and Py04_0364) of these Pyrococcus LysX-like proteins do not contain a tyrosine residue at strands β10 or β11, as observed in LysX/ArgX and ArgX, respectively. Because these lysX homologs are linked to argG and argH homologs, they presumably function as argX. Other LysX proteins from Pyrococcus have a threonine residue at the large loop, corresponding to Thr196 in TkLysX/ArgX; therefore, this threonine may be the factor affecting whether it has the ability to act as ArgX or LysX. Taken together with PNA2_0499, which has different residues at the signature motif, we speculate that these LysX family proteins, the gene of which are clustered with the arginine biosynthetic genes, are still at the halfway point of evolution to ArgX for the biosynthesis of arginine.

The evolutionary relationship among lysine, arginine, and leucine biosynthetic pathways has already been discussed (9, 20, 21), and it has been suggested that these pathways have evolved from a common ancestral multifunctional pathway. This study provides experimental evidence for the existence of an ancestral-type multifunctional metabolic pathway composed of enzymes with wide substrate specificities, as predicted from the patchwork hypothesis. There are two major known biosynthetic pathways for lysine, the DAP and AAA pathways (8, 22). Several enzymes in the DAP pathway (products of the ask, asd, dapC, and dapE genes) share amino acid sequences with those in the LysW-mediated AAA pathway: LysZ, LysY, LysJ, and LysK, respectively. This result suggests that these two pathways, which are considered to have different origins, are also evolutionarily related. The mechanism underlying substrate recognition and information on the molecular evolution of LysX family proteins revealed in the present study will contribute to a better understanding of the evolution of biosynthetic pathways in ancient organisms.

Experimental Procedures

Preparation of Expression Vectors—Regarding the heterologous expression of lysine biosynthetic genes in E. coli, we prepared vectors that directed the production of gene products with C- or N-terminal His tags. We also constructed vectors that expressed TK0278 and TK0279 without tags for crystallization. To amplify each gene, a polymerase chain reaction (PCR) was performed with appropriate primers (Table 2) using the genome of T. kodakarensis, and the amplified DNA fragments were cloned into pBlueScriptII KS(+) (Agilent Technology) to verify their sequences. Most amplified DNA fragments were inserted into the pET26b(+) (Novagen) vector, whereas

T. paralvinellae

T. kodakarensis

FIGURE 10. Specific activity of TkLysX/ArgX mutants. Bars filled with white and black indicate activity toward AAA and glutamate, respectively. Enzyme assays were performed in triplicate. Standard deviations were shown for each bar.

FIGURE 11. Lysine and arginine biosynthetic gene cluster in T. paralvinellae and T. kodakarensis.
the fragment of TK0274 was introduced into the pHIS8 vector (23) for expression. The constructed plasmids were named pET26b-TK0279, pET26b-TK0278Chis, pET26b-TK0278, pET26b-TK0277Chis, pET26b-TK0275Chis, and pET26b-TK0275Chis, and pHIS8-TK0274.

Preparation of Proteins—The recombinant proteins TK0279 (TkLysW), TK0278 (TkLysX/ArgX) (with and without the (His)6 tag), TK0276, and TK0275 were produced in E. coli BL21-CodonPlus (DE3)-RIL (Agilent Technologies). To increase the amount of soluble proteins, we introduced the expression plasmids for TK0277 and TK0274 into E. coli BL21(DE3) harboring pT-groE (24). These transformants were grown in 2× YT (5 g/liter of NaCl, 10 g/liter of Bacto-yeast extract, and 16 g/liter of Bacto-tryptone) broth supplemented with 30 μg/ml of chloramphenicol and 50 μg/ml of kanamycin at 37 °C for 3 h. Gene expression was then induced by the addition of isopropyl β-D-thiogalactoside (final concentration of 1 mM) for TkLysW, TK0277, and TK0274 and cultured at 37 °C for an additional 12–14 h. In the expression of TkLysW/ArgX, TK0276 (TkLysX/ArgB), and TK0275, we added at a final concentration of 0.1 mM isopropyl β-D-thiogalactoside to the culture after a 3-h incubation at 37 °C and the culture was continued at 25 °C for an additional 12–14 h.

The enzymes in the activity assay and in vitro reconstruction were purified in their His-tagged forms as follows. Harvested cells were washed with buffer A (20 mM HEPES–NaOH, pH 8.0) and sonicated in buffer B (20 mM HEPES–NaOH, pH 8.0, 150 mM NaCl) for TK0276 and TK0275, and in buffer C (20 mM HEPES–NaOH, pH 8.0, 300 mM NaCl, 10% glycerol) for TkLysX/ArgX, TK0277, and TK0274. The supernatants obtained from the centrifugation of these sonicates were heated at 80 °C for 20 min. The heat-stable fractions were applied to a nickel–nitrilotriacetic acid column (Novagen) equilibrated with buffers B or C supplemented with 20 mM imidazole, pH 8.0, and washed with the same buffer. Each protein was eluted with buffers B or C supplemented with 200 mM imidazole, pH 8.0.

We used TkLysW and TkLysX/ArgX, both without (His)6 tags, for crystallization. The purification of TkLysW was performed in the same manner as that for TkLysW (12). TkLysW/ArgX was purified as follows. Cells harboring pET26b-TK0278 were harvested and washed with buffer D (20 mM Tris–HCl, pH 8.0). They were then resuspended in buffer E (20 mM Tris–HCl, pH 8.0, 150 mM NaCl) and disrupted by sonication. The supernatant obtained from the cell lysate was heated at 80 °C for 20 min and heat-stable proteins containing TkLysX/ArgX were precipitated with 65% ammonium sulfate. The precipitate was dissolved in buffer F (20 mM Tris–HCl, pH 8.0, 1 M ammonium sulfate) and applied to HiLoad 16/10 phenyl-Sepharose HP (GE Healthcare). Proteins were eluted by a linear gradient of buffers F to D. Fractions containing TkLysX/ArgX were concentrated using VIVASPIN (MWCO 10,0000) (Sartorius), and applied to a ResourceQ column (GE Healthcare) equilibrated with buffer A. TkLysX/ArgX was eluted by a linear gradient of buffers D to G (20 mM Tris–HCl, pH 8.0, 1.0 M NaCl). TkLysX/ArgX was further applied to gel filtration chromatography with HiLoad 26/60 Superdex 75 pg (GE Healthcare) with buffer E, and used as a purified sample for crystallization.

The purities of proteins, except for TkLysW, were confirmed using 12% SDS-PAGE, purified proteins were concentrated by VIVASPIN (MWCO 10,000), and their concentrations were measured using a Protein Assay kit (Bio-Rad). Regarding TkLysW, we checked its purity using 12% Tricine SDS-PAGE (25), concentrated it using VIVASPIN with MWCO 3,000, and measured its concentration with a BCA Protein Assay kit (Pierce).

Activity Assay of TK0278 and TK0276—The LysX and ArgX activities of TkLysW/ArgX were measured by monitoring inorganic phosphate released during the reaction by Determinar L IP (Kyowa Medex) as described previously (5). AAA and glutamate at 5 mM were used as substrates. The activity assays for the TkLysX/ArgX mutants were performed in the same manner with a fixed concentration of TK0278 at 0.5 mg/ml.

### Table 2

Oligonucleotides used in this study

| Name | Nucleotide sequence |
|------|---------------------|
| pET26b-TK0279-Fw | 5'-CTCTAGACAGATGTTGGAATGCCTCTC-3' |
| pET26b-TK0279-Rv | 5'-GAATTCTGACCCCGATGCCTCTC-3' |
| pET26b-TK0278-Fw | 5'-GAATTCTGACGCTGGCCGCTCTC-3' |
| pET26b-TK0278-Rv | 5'-GTCTAGACGCTGGCCGCTCTC-3' |
| pET26b-TK0277-Fw | 5'-GAATTCTGACGCTGGCCGCTCTC-3' |
| pET26b-TK0277-Rv | 5'-GTCTAGACGCTGGCCGCTCTC-3' |
| pET26b-TK0276-Fw | 5'-GAATTCTGACGCTGGCCGCTCTC-3' |
| pET26b-TK0276-Rv | 5'-GTCTAGACGCTGGCCGCTCTC-3' |
| pET26b-TK0275-Fw | 5'-GAATTCTGACGCTGGCCGCTCTC-3' |
| pET26b-TK0275-Rv | 5'-GAATTCTGACGCTGGCCGCTCTC-3' |
| pHIS8-TK0274-Fw | 5'-CCCAAGCTTGACTGAATGATATCTGCTATCCGAC-3' |
| pHIS8-TK0274-Rv | 5'-CCCAAGCTTGACTGAATGATATCTGCTATCCGAC-3' |
| TK0279E42R-Fw | 5'-GCTCTAGACATATGCCGCTCTACAGGAAG-3' |
| TK0279E42R-Rv | 5'-GCTCTAGACATATGCCGCTCTACAGGAAG-3' |
| TK0278Y175F-Fw | 5'-GTGAACCCTGGGAACCTGGCTACTC-3' |
| TK0278Y175F-Rv | 5'-GTGAACCCTGGGAACCTGGCTACTC-3' |
| TK0278N250GA251F-Fw | 5'-GACAGAATTCGACGCTGGCCGCTCTC-3' |
| TK0278N250GA251F-Rv | 5'-GACAGAATTCGACGCTGGCCGCTCTC-3' |
| TK0278I185Y-Fw | 5'-GACACATATGCTCCCTAAATTTAGGGAAG-3' |
| TK0278I185Y-Rv | 5'-GACACATATGCTCCCTAAATTTAGGGAAG-3' |
| TK0278N250GA251S-Rv | 5'-GACATAGAATTCGACGCTGGCCGCTCTC-3' |
| TK0278N250GA251S-Fw | 5'-GACATAGAATTCGACGCTGGCCGCTCTC-3' |
| TK0278E42R-Fw | 5'-GCTCTAGACATATGCCGCTCTACAGGAAG-3' |
| TK0278E42R-Rv | 5'-GCTCTAGACATATGCCGCTCTACAGGAAG-3' |
Lysine Biosynthesis of *T. kodakarensis*

TkLysZ/ArgB activity was analyzed as follows. Because wild-type TkLysW has no cleavage sites (lysine or arginine residue) by trypsin, we introduced the E42R mutation using appropriate primers (Table 2), which facilitated the analysis of the structure of the *C* terminus of TkLysW by LC-MS/MS. TkLysW(E42R) was purified with a heat treatment and successive DE52 anion exchange column (Whatman). Purified TkLysW(E42R) was treated with TkLysX/ArgX to yield TkLysW-γ-AAA and TkLysW-γ-Glu in the same manner to prepare TkLysW-γ-AAA as reported previously (12). The TkLysZ/ArgB reaction was performed in the same manner as described previously in the presence of hydroxylamine. The reaction mixture was applied to Tricine SDS-PAGE and we analyzed TkLysW derivatives in more detail using MALDI-TOF-MS. In-gel digestion and subsequent LC-MS/MS and MALDI-TOF-MS analyses were performed as described previously (5).

**In Vitro Reconstruction of Lysine and Ornithine Biosynthesis**—In the *in vitro* synthesis of lysine and ornithine from AAA and glutamate, respectively, we prepared recombinant proteins of TkLysW, TkLysX/ArgX, TkLysZ/ArgB, TK0277, TK0275, and TK0274 as described above. The reaction mixture contained 100 mM HEPES-NaOH, pH 8.0, 330 μM TkLysW, 10 mM AAA, or glutamate, 10 mM ATP, 1 mM MgSO₄, 2.5 mM NADPH, 0.1 mM PLP, 0.1 mM ZnSO₄, and 0.1 mM CoCl₂. The reaction was initiated by adding each enzyme to the reaction mixture at a concentration of 0.5 mg/ml and incubating at 60°C for 1 h. We also performed similar reactions in mixtures lacking each protein. After being incubated, an equal amount of 6% trichloroacetic acid was added to precipitate the proteins. Amino acids in the supernatant were analyzed with the High-speed Amino Acid Analyzer L-8900 (Hitachi High-Tech).

**Crystallization, Data Collection, and Processing**—To obtain the crystal structure of the TkLysX/ArgX-TkLysW complex, we screened crystallization conditions. Purified TkLysX/ArgX (7.5 mg/ml) and TkLysW (2.5 mg/ml) were mixed with 10 mM AMP-PNP, 10 mM AAA, or 50 mM glutamate and 10 mM MgSO₄. Screening and subsequent optimization were performed using several crystallization screening kits by the hanging drop vapor diffusion method at 20°C. Crystals were obtained from the condition containing 20% (v/v) 2-methyl-2,4-pentanediol, 6.6% (w/v) polyethylene glycol (PEG)-3350, 0.066 M imidazole, pH 6.5, and 0.132 M ammonium sulfate, which was prepared by dilution (66%) of No. 29 in the crystallization screening kit, Precipitant Synergy P64 (Rigaku Regents).

A crystal was directly flash-cooled by a nitrogen gas stream at 95 K. Diffraction data were collected at 1.000 Å with a CCD camera (ADSC Quantum 280) on beamline NE3A of the Photon Factory, High Energy Accelerator Research Organization (KEK) (Tsukuba, Japan). Diffraction data were indexed, integrated, and scaled using the HKL2000 program suite (26) and the crystal belonged to the space group *P*2₁ with unit cell parameters of *a* = 76.2 Å, *b* = 135.1 Å, *c* = 137.0 Å, *β* = 97.5°.

**Structural Determination and Refinement**—Structures were determined by the molecular replacement method using the TtLysX (PDB code 3VPD) and TtLysW-γ-AAA (PDB code 3WWL) structures as searching models. Molecular replacement was conducted by the program Phaser (27) in the CCP4 program suite (28) and subsequent structural refinement was performed using Refmac5 (29) and Coot (30). The overall geometry of the structural model was checked by MolProbity (31), and the figures were prepared using PyMOL (32). Data collection and refinement statistics are summarized in Table 1. The refined structure of TkLysX/ArgX-TkLysW-γ-AAA was deposited to the Protein Data Bank with accession number 5K2M. The glutamate binding model structure of TkLysX/ArgX was created as described below. Glutamate was placed manually in the active site by superposing it with the glutamate adduct attached to StLysW-γ-Glu of the StArgX pseudo-Michaelis complex. The initial model structure was then refined by energy minimization using the CFF force field in the Discover Studio (BIOVIA).

**Preparation of TkLysX/ArgX Mutants**—The expression plasmids for the TkLysX/ArgX mutants used in the activity assay were prepared by site-directed mutagenesis using QuikChange site-directed mutagenesis kits (Agilent Technology) for pET26b-TK0278CHis with appropriate primers (Table 2). The purification of these mutants was performed in the same manner for wild-type TkLysX/ArgX.

**Bioinformatic Analysis of LysX Family Proteins**—In the multiple amino acid sequence alignment, we selected LysX family proteins from the KEGG database, particularly among organisms belonging to Thermococcales. The multiple amino acid sequence alignment was analyzed by Clustal W (33) and depicted using ESPript (34).

**Author Contributions**—A. Y. performed all the experiments and wrote the manuscript; T. T. performed the crystallographic analysis; H. A. and T. K. planned the experiments; M. N. planned the experiments, analyzed data, and wrote the manuscript.

**Acknowledgment**—We are grateful to the staff at the Photon Factory for their assistance with data collection, which was approved by Photon Factory Program Advisory Committee Proposals 2011G525, 2013G619, and 2015G552.

**References**

1. Jensen, R. A. (1976) Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* 30, 409–425.

2. Lazcano, A., and Miller, S. L. (1996) The origin and early evolution of life: prebiotic chemistry, the pre-RNA world, and time. *Cell* **85**, 793–798.

3. Alifano, P., Fanì, R., Liò, P., Lazcano, A., Bazzicalupo, M., Carloni, M. S., and Bruní, C. B. (1996) Histidine biosynthetic pathway and genes: structure, regulation, and evolution. *Microbiol. Rev.* **60**, 44–69.

4. Copley, R. R., and Bork, P. (2000) Homology among (βα)₈ barrels: implications for the evolution of metabolic pathways. *J. Mol. Biol.* **303**, 627–641.

5. Horie, A., Tomita, T., Saiki, A., Kono, H., Taka, H., Mineki, R., Fujimura, T., Nishiyama, C., Kuzuyama, T., and Nishiyama, M. (2009) Discovery of proteinaceous N-modification in lysine biosynthesis of *Thermus thermo-philus*. *Nat. Chem. Biol.* **5**, 673–679.

6. Kobashi, N., Nishiyama, M., and Tanokura, M. (1999) Aspartate kinase-independent lysine synthesis in an extremophilic bacterium, *Thermus thermophilus*: lysine is synthesized via α-aminoadipic acid not via diaminopimelic acid. *J. Bacteriol.* **181**, 1713–1718.

7. Nishida, H., Nishiyama, M., Kobashi, N., Kosuge, T., Hoshino, T., and Yamane, H. (1999) A prokaryotic gene cluster involved in synthesis of...
lysin through the amino adipate pathway: a key to the evolution of amino acid biosynthesis. *Genome Res.* **9**, 1175–1183
8. Umbarger, H. (1978) Amino acid biosynthesis and its regulation. *Annu. Rev. Biochem.* **47**, 532–606
9. Irvin, S. D., and Bhattacharjee, J. K. (1998) A unique fungal lysine biosynthesis enzyme shares a common ancestor with tricarboxylic acid cycle and leucine biosynthetic enzymes found in diverse organisms. *J. Mol. Evol.* **46**, 401–408
10. Velasco, A. M., Leguina, J. I., and Lazzano, A. (2002) Molecular evolution of the lysine biosynthetic pathways. *J. Mol. Evol.* **55**, 445–459
11. Ouchi, T., Tomita, T., Horie, A., Yoshida, A., Takahashi, K., Nishida, H., Lassak, K., Taka, H., Mineki, R., Fujimura, T., Kosono, S., Nishiyama, C., Masui, R., Karukitsu, S., Albers, S. V., Kuzuyama, T., and Nishiyama, M. (2013) Lysine and arginine biosyntheses mediated by a common carrier protein in *Sulfolobus* sp. *Nat. Chem. Biol.* **9**, 277–283
12. Yoshida, A., Tomita, T., Fujimura, T., Nishiyama, C., Kuzuyama, T., and Nishiyama, M. (2015) Structural insight into amino group-carrier protein-mediated lysine biosynthesis: crystal structure of the LysZ/LysW complex from *Thermus thermophilus*. *J. Biol. Chem.* **290**, 435–447
13. Leisinger, T., and Haas, D. (1975) N-Acetylglutamate synthase of *Escherichia coli* regulation of synthesis and activity by arginine. *J. Biol. Chem.** 250**, 1690–1693
14. Baetens, M., Legrain, C., Boyen, A., and Glansdorff, N. (1998) Genes and enzymes of the acetyl cycle of arginine biosynthesis in the extreme thermophilic bacterium *Thermus thermophilus* HB27. *Microbiology* **144**, 479 –492
15. Martin, P. R., and Mulks, M. H. (1992) Sequence analysis and complementation studies of the *argF* gene encoding ornithine acetyltransferase from *Neisseria gonorrhoeae*. *J. Bacteriol.* **174**, 2694–2701
16. Hara, T., Kato, H., Katsube, Y., and Oda, J. (1996) A pseudo-Michaelis quaternary complex in the reverse reaction of a ligase: structure of *Escherichia coli* B glutathione synthetase complexed with ADF, glutathione, and sulfate at 2.0 Å resolution. *Biochemistry* **35**, 11967–11974
17. Tavlik, D. S. (2010) Messy biology and the origins of evolutionary innovations. *Nat. Chem. Biol.* **6**, 692–696
18. Fukui, T., Atomi, H., Kanai, T., Matsumi, R., Fujiwara, S., and Imanaka, T. (2005) Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1 and comparison with *Pyrococxus* genomes. *Genome Res.* **15**, 352–363
19. Jung, J. H., Kim, Y. T., Jeon, E. J., Seo, D. H., Hensley, S. A., Holden, J. F., Lee, J. H., and Park, C. S. (2014) Complete genome sequence of hyperthermophilic archaeon *Thermococcus* sp. ES1. *J. Biotechnol.* **174**, 14–15
20. Fondi, M., Brilli, M., Emiliani, G., Paffetti, D., and Fani, R. (2007) The primordial metabolism: an ancestral interaction between leucine, arginine, and lysine biosynthesis. *J. Biol. Chem.* **282**, 479–492
18. Fukui, T., Atomi, H., Kanai, T., Matsumi, R., Fujiwara, S., and Imanaka, T. (2007) Surprising arginine biosynthesis: a reappraisal of the enzymology and evolution of the pathway in microorganisms. *Microbiol. Mol. Biol. Rev.* **71**, 36–47
22. Vogel, H. (1964) Distribution of lysine pathways among fungi: evolutionary implications. *Am. Nat.* **98**, 435–446
23. Jez, J. M., Ferrer, J. L., Bowman, M. E., Dixon, R. A., and Noel, J. P. (2000) Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. *Biochemistry* **39**, 890–902
24. Yasukawa, T., Kamei-Ishii, C., Maekawa, T., Fujimoto, I., Yamamoto, T., and Ishii, S. (1995) Increase of solubility of foreign proteins in *Escherichia coli* by coproduction of the bacterial thioredoxin. *J. Biol. Chem.* **270**, 25328–25331
25. Schägger, H. (2006) Tricine-SDS-PAGE. *Nat. Protoc.* **1**, 16–22
26. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326
27. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
28. Collaborative Computational Project, Number 4 (1994) The CCP4 Suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763
29. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255
30. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
31. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–383
32. Delano, W. L. (2002) The PyMol Molecular Graphics System, Version 1.6.0.0, Schrödinger LLC, San Diego
33. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680
34. Gouet, P., Courcelle, E., Stuart, D., and Mézot, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308