Activation reaction:



FIG S1. The two-step reaction catalyzed by lipoate-protein ligase (Lpl). Abbreviations: PPi, pyrophosphate; P, lipoate-dependent protein.



FIG S2. SDS-PAGE analysis of the purified TK1234 (A) and TK1908 (B) recombinant proteins. Four micrograms each of purified TK1234 and TK1908 recombinant proteins were applied. Gels were stained with Coomassie Brilliant Blue. M indicates molecular mass marker.



FIG S3. Association analysis of TK1234 and TK1908 proteins. (A) Chromatogram of gel filtration chromatography analyzing TK1234 protein (100 μ M), TK1908 protein (100 μ M), and their mixture. A Superdex 200 Increase 10/300 GL column was used. The mixture was composed of equimolar TK1908 protein and TK1234 protein (100 μ M each). (B) SDS-PAGE analysis of a fraction corresponding to the peak P1 in the protein mixture solution (15.0 ml to 15.5 ml). M indicates the molecular mass marker.



FIG S4. Progress curve of lipoyl-peptide (LP) formation catalyzed by single TK1908 protein. The reaction mixture was composed of 50 mM HEPES (pH 7.5 at 85°C), 2 mM MgCl₂, 10 mM lipoate, 100 μ M synthetic peptide, 5 mM ATP, and 0.7 μ g/ μ l TK1908 protein.



FIG S5. Identification of ligase reaction products by LC-MS analysis. The reactions were carried out with octanoate (A), heptanoate (B), hexanoate (C), pentanoate (D) and butyrate (E) as substrates and their reaction products were individually examined. Standard modified peptides were chemically synthesized. The chromatograms of the compounds whose exact masses corresponded to those of the modified peptides are shown. U1-U5 indicate the by-products with the same exact masses to those of modified peptides. Abbreviations: OP, standard octanoyl-peptide; HP, standard hexanoyl-peptide; BP, standard butyryl-peptide; AU, arbitrary units.





FIG S6. Lipoyl synthase activity measurement with HPLC. The reactions were carried out with octanoyl-peptide (A), hexanoyl-peptide (B), butyryl-peptide (C), and free octanoate (supplemented with or without peptide) (D) as substrates and their reaction products were individually examined. Standard lipoyl-peptide, standard octanoyl-peptide, standard hexanoyl-peptide, and standard butyryl-peptide were chemically synthesized. Abbreviations: LP, lipoyl-peptide; RLP, reduced lipoyl-peptide; OP, octanoyl-peptide; HP, hexanoyl-peptide; BP, butyryl-peptide; LA, lipoate; RLA, reduced lipoate; U1-U4, unidentified compounds; AU, arbitrary units.



FIG S7. Construction of gene disruption mutants. (A) Schematic drawing for gene disruption in *T. kodakarensis*. As an example, the disruption of TK1908 is shown. *T. kodakarensis* KU216 (Δ*pyrF*) was used as the host strain. Pop-in/pop-out strategy using a *pyrF* gene as a selectable marker was applied. The gene disruption strain Δ TK1234 was constructed with the same method. (B-E) Genotypic analyses of the gene disruption strains. PCR analyses of the TK1234 and TK1908 gene disruption strains (Δ TK1234 and Δ TK1908) were carried out with primer sets that anneal outside of the TK1234- and TK1908-flanking regions for homologous recombination (B and C), respectively, and within the TK1234 and TK1908 genes (D and E), respectively. Abbreviations: M1, λ -Hind III marker (λ -phage DNA digested with HindIII); M2, 20 bp DNA Ladder (Dye Plus) (Takara Bio); M3, 100 bp DNA Ladder (Dye Plus) (Takara Bio); WT, *T. kodakarensis* wild-type KOD1; H, the host strain KU216; D1, Δ TK1234; D2, Δ TK1908.

	β1	α1		β2	β3	$\alpha 2$
Ta-LplA-N	>	22222222222222	тт	>	—→ тт	lll
Ta-LplA-N 1 Tk-Lpl-N 1 Ec-LplA 1 Bs-LplJ 1 Mhp-Lpl 1 Lmo-LplA 1 Sp-LplA 1 Ef-LplA1 1	MEGRLLLLETPG. MRFIPLIVARP. MSTLRLLISDS. MLFIDNQNIN MYLIEPKRNGKWV MYFIDNNNEK .MKYIINHSN	NTRMSLAYDEATYRSF EVQMAIDEATMRAR YDPWFNLAVEECIFRQM DPRINLAIEYCVKHL FDGAILLAIQYWAIKNL DPRINLAVEEFILTEL DTAFNIALEYAFKHL DTAFNIALEYAFKHL	QYGDKP IEGKVPDI PATQF DPEQ(KLDEI NLDEI LDED(LDED(ILRFYRHDE VRLYAFSPS VLFLWRNAI VLFYVNQE IVFPYICDE VLLFYINKE IFLLWINKE	RSVIIGYFQ SSVIIGRFQ TVVIGRAQ SIIIGKNQ SIIIGYFQ SIIIGRNQ SIIVGRHQ SIIVGRHQ	VAEEE SVVHD NPWKE NTIEE NPSVE NTVEE NTIEE NTIEE

		α3	β4	β5	β6	α4
Ta-LplA-N		ellee	>	— тт		ellellele
Ta-LplA-N	57	VDLDYMKKNG	IMLA <mark>RR</mark> YT	<mark>GGG</mark> AVYH <mark>D</mark> LG	DLNFSVVRSSDD	.MDITSMFRTMNEAV
Tk-Lpl-N	56	VNLDEAKKLR	IPVVRRIT	GGGSVFH <mark>D</mark> EYG.	EITYSIVIGEDYHPA	LKNVEESYRYLAGPL
Ec-LplA	56	CNTRRMEEDN	VRLA <mark>RR</mark> SS	GGGAVFH <mark>D</mark> LG	NTCFTFMAGKPE	YDKTISTSIV
Bs-LplJ	54	INTKYVEENG	IIVV <mark>RR</mark> LS	GGGAVYH <mark>D</mark> LG	NLNFSFITKDDG	DSFHNFKKFTEPV
Mhp-Lpl	58	VNLELLKQKN	IEVVRRDT	<mark>GGG</mark> AIYL <mark>D</mark> RNGV	NFCFSFPYEKNK	NLLGNYAQFYDPV
Lmo-LplA	54	IDTEYVEKND	VIVV <mark>RR</mark> LS	GGGAVYH <mark>D</mark> EG	NLNFSFITEDDG	ESFHNFAKFTQPI
Sp-Lp1A	53	INRDYVRENG	IEVVRRIS	GGGAVYH <mark>D</mark> LN	NLNYTIISKEDE	NKAFDFKSFSTPV
Ef-LplA1	53	INRDYVRENG	IEVVRRIS	<mark>GGG</mark> AVYH <mark>D</mark> LN	NLNYTIISKEDE	NKAFDFKSFSTPV

			_{в7} Са	apping lo	ор	β8	β9	β10
Ta-LplA-N		وووووو		тт	тт —		<u> </u>	` >
Ta-LplA-N	111	VNSLRILGLI	ARPGELN	DVSIPVN	KKTDI	MAGEK <mark>K</mark> I	M <mark>G</mark> A <mark>A</mark> GAMR	KGAKLWHAAMLVHTD
Tk-Lpl-N	115	VDALKDLGLF	. AGF <mark>SGL</mark> N		DI	VANGK <mark>K</mark> I	S <mark>G</mark> S <mark>A</mark> QTRR	RGVILQHGTFMYATR
Ec-LplA	106	LNA <mark>L</mark> NAL G VS	. A E A S G R <mark>N</mark>	D	LVVKT	VEGDR <mark>K</mark> V	′S <mark>G</mark> S <mark>A</mark> YRET:	K D R G F H H G T L L L N A D
Bs-LplJ	107	IQALHQLGVE	. A E L S G R <mark>N</mark>		DI	VVDGR <mark>K</mark> I	SGNAQFAT	KGRIFSHGTLMFDSA
Mhp-Lpl	113	IKVLQNIGIF	KNVQFSGK <mark>N</mark>		DL	QIEGKK	'S <mark>G</mark> A <mark>A</mark> MSLV:	NDRIYAGFSLLYDVD
Lmo-LplA	107	VEALKRLGVN	I. AELKGR <mark>N</mark>		DL	LIDGFK	′S <mark>G</mark> N <mark>A</mark> QFAT:	KGKMFSHGTLMYDLN
Sp-LplA	106	INTLAQLGVF	. A E F T G R <mark>N</mark>		DL	EIDGK <mark>K</mark> F	C <mark>GNA</mark> QAYI	NGRIMHHGCLLFDVD
Ef-LplA1	106	INTLAQLGVF	. A E F T G R <mark>N</mark>		D L	EIDGKK	C <mark>GNA</mark> QAYI	NGRIMHHGCLLFDVD

		α5		η1	η2	α6
Ta-LplA-N		eeeeee	•	عفع	ee.e	2 $2222222222222222222222222222222$
Ta-LplA-N	170	LDMLSAVLK	V P D E <mark>K</mark> F R D <mark>K</mark> I A F	K <mark>S</mark> TRE RV 2	ANVT.DFVD	VSIDEVRNALIRGFSETLH
Tk-Lpl-N	164	VEILGRVLR	V S <mark>K E K</mark> L K D <mark>K</mark> G I S	S <mark>S</mark> IWE <mark>RV</mark> :	「 <mark>TL</mark> EREGVト	(LSRWEAYELLKESFSNA
Ec-LplA	159	LSRLANYLNI	P D K K <mark>K</mark> L A A <mark>K</mark> G I 1	[<mark>S</mark> VRS <mark>RV</mark>]	INLTELLPO	JITHEQVCEAITEAFFAHYG
Bs-LplJ	156	IDHVVSALK	V K K D <mark>K</mark> I E S <mark>K</mark> G I F	K <mark>S</mark> IRS <mark>RV</mark> A	ANISEFLDI)KMTTEEFRSHLLRHIFNTNDV
Mhp-Lpl	163	FDFIGKILT	P N Q K <mark>K</mark> I E A <mark>K</mark> G I F	(SVSQRV)	[N L K N K L S F	<pre>KEYQNFSIFEIKDLFLTEFLKVNSV</pre>
Lmo-LplA	156	LDNVAASLKI	P R K D <mark>K</mark> I E S <mark>K</mark> G I F	K <mark>s</mark> vrs rv z	ANISDFMDÇ	2EMTTEEFRDLLLLYIFGVEKV
Sp-LplA	155	LSVLANALK	V S K D <mark>K</mark> F E S <mark>K</mark> G V F	(SVRA <mark>RV</mark>)	[NIINELPF	KITVEKFRDLLLEYMKKEY
Ef-LplA1	155	LSVLANALK	V S <mark>K D K</mark> F E S <mark>K </mark> G V F	(SVRA <mark>RV</mark>)	[<mark>ni</mark> ine l př	KITVEKFRDLLLEYMKKEY

		β11	α	7	α8		
Ta-LplA-N			فعفعفعه	۹۵	eeee		
Ta-LplA-N	224	IDFREDTI	TEKEESLAR	ELFDKK	YSTEEWNMGL	LRKEVV	
Tk-Lpl-N	217	FELEEGEL	TDYELELAE	K <mark>L</mark> IEE <mark>K</mark>	K Y R N P K W N E M R		
Ec-LplA	214	ERVEAEII	SPNKTPDLP	NFAETFA <mark>F</mark>	RQSSWEWNFGQ	APAFSHLLDE	R F T W G G V E L H F D V E K
Bs-LplJ	213	GNVPEYKL	TEKDWETIH	QISKEF	RYQNWDWNYGR	SPKFNLNHSK	RYPVGSIDLHLEVKK
Mhp-Lpl	223	EKFKKYEL	TDSDWVQID	KMVAEK	KYKNWDFVWGL	SPNYSFNRSI	RTKVGTITFSLEINE
Lmo-LplA	213	EDVKEYKL	TAADWEKIH	EISAK <mark>F</mark>	R Y G N W D W N Y G K	SPKFDLTRTK	R F P V G A V D V R L N V Q K
Sp-LplA	210	PEMTEYVF	SEEELAEIN	RIKDT <mark></mark>	K F G T W D W N Y G K	SPEFNVRRGI	K F T S G K V E V F A N V T E
Ef-LplA1	210	PEMTEYVF	SEEELAEIN	RIKDTK	KFGTWDWNYGK	SPEFNVRRGI	KFTSGKVEVFANVTE

FIG S8. Amino acid sequence alignment of the large domains of LpIs. The sequences of *Tk*-LpI-N (TK1908) and the large domains of *T. acidophilum* LpI (*Ta*-LpIA-N), *E. coli* LpI (*Ec*-LpIA), *B. subtilis* LpI (*Bs*-LpIJ), *M. hyopneumoniae* LpI (*Mhp*-LpI), *L. monocytogenes* LpI (*Lmo*-LpIA), *Streptococcus pneumoniae* LpI (*Sp*-LpIA), and *Enterococcus faecalis* LpI (*Ef*-LpIA1) were aligned. The secondary structure elements of *Ta*-LpIA-N (PDB entry 2ARS) are shown above the sequence (squiggles, α -helices; arrows, β -strands; TT, β -turns; η , 3/10-helices). The RRX(T/S)GGG, KX₂GXA, and KX₃KX₃SX₃RV motifs conserved among LpI members are indicated by green, magenta and gray bars, respectively. The capping loop motif of *Ta*-LpIA-N (residues 124-137) is underlined in yellow. Similar residues in the columns are shown by red letters, and identical residues are shown by white letters on a red background. The columns that display high similarity (global score higher than 0.7) are framed in blue. The dots above the alignment indicate every 10th residue of the top sequence.

Ta-LplA-C	1	
Tk-Lpl-C	1	
Ec-LplA	230	PNFAETFARQS <mark>SWEW</mark> NFGQAPAFSHLLDERFTW <mark>G</mark> GVELHFDVEKGHITRAQVFT <mark>D</mark> SLNPA
Bs-LplJ	227	ſĬĦQĬŚĸĔŖŶQ <mark>ŊŴŊŴŊ</mark> ŶĠŖŚ₽ĸĔŊĹŊĦŚĸŖŶŶŶ <mark>Ġ</mark> ŚĬŊĹĦĹĘŴĸĸĠĸĬĔŊĊĸĬĔĠ <mark>Ď</mark> ĔĔĠŸĠ
Mhp-Lpl	237) I D KMVAEKYK <mark>N W DFVW GLSPNYSFNR SIR TKV<mark>G</mark>TITFSLEINEGKISKIKISG<mark>D</mark>FFPKK</mark>
Lmo-LplA	227	KIHE I SAKRYGNWDWNYGKSPKFDLTRTKRFPV <mark>G</mark> AVDVRLNVOKGVITDIKIFG <mark>D</mark> FFGVK
Sp-LplA	224	EINRIKDTKFGTWDWNYGKSPEFNVRRGIKFTS <mark>G</mark> KVEVFANVTESKIODIKIYG <mark>D</mark> FFGIE
Ef-LplA1	224	SINRIKDTKFGTWDWNYGKSPEFNVRRGIKFTS G KVEVFANVTESKIQDIKIYG D FFGIE
-		
Ta-LplA-C	47	SDSINRLEDMERGSSIEKINDIIRDFYNOGVITPGVEPEDFIOALRVI
Tk-Lpl-C	40	E E TVHELE S R <mark>L</mark> EGHKLEE LEG I I D E F F AMRLD V E M P Y I N V E D F K I A L K K A L E G
Ec-LplA	290	2. LEALAGR <mark>L</mark> OGCLYRADMLOOECEALLVDFPEOEKELRELSAWMAGAVR
Bs-LplJ	287	D VSEIENL <mark>I</mark> VGKOYERSVIÄDVLEGVNLKHYFGN.ITKEDFLDLIY
Mhp-Lpl	297	LLELENFLIGTKLTODOLLNRLKDAKLDDYFSOKIDEEEICNLLLN
Lmo-LplA	287	N. VADIEEKTVNTTYKŘEVLAEALVDIDVKEYFGŇ, ITKDEFLDLLY
Sp-LplA	284	
Ef-LplA1	284	

FIG S9. Amino acid sequence alignment of the small domains of Lpls. The sequences of *Tk*-Lpl-C and the small domains of *T. acidophilum* Lpl (*Ta*-LplA-C), *E. coli* Lpl (*Ec*-LplA), *B. subtilis* Lpl (*Bs*-LplJ), *M. hyopneumoniae* Lpl (*Mhp*-Lpl), *L. monocytogenes* Lpl (*Lmo*-LplA), *S. pneumoniae* Lpl (*Sp*-LplA), and *E. faecalis* Lpl (*Ef*-LplA1) were aligned. Residues are colored and indicated as described in Fig. S8.



FIG S10. Structural superimpositions of *Tk*-Lpl-N with Lpls from other species. Structural superimpositions of *Tk*-Lpl-N (cyan) with the large domain of *T. acidophilum* Lpl (*Ta*-LplA-N, PDB entry 2ARS, gray) (A), *T. acidophilum* Lpl (*Ta*-LplA, PDB entry 3R07, gray) (B), *M. hyopneumoniae* Lpl (*Mhp*-Lpl, PDB entry 6JOM, light blue) (C), and *E. coli* Lpl (*Ec*-LplA, PDB entry 1X2G, dark green) (D). Only the large domains are shown here. The conserved motifs RRX(T/S)GGG and KX₂GXA in the active site of *Tk*-Lpl-N structure are indicated in green and magenta, respectively. The capping loop motif and the corresponding regions of *Tk*-Lpl-N (red), *Ta*-LplA-N (yellow), *Ta*-LplA (yellow), *Mhp*-Lpl (blue), and *Ec*-LplA (blue) are colored.



FIG S11. Octanoyl transferase activity measurement with HPLC. The octanoyl transferase activity was determined based on a previous study (1). The reaction mixture (50 μ l) was composed of 50 mM HEPES (pH 7.5 at 85°C), 100 mM KCl, 5 mM DTT, 100 μ M synthetic peptide, 100 μ M octanoyl-coenzyme A (CoA), 5 μ g TK1908 protein (and 5 μ g TK1234 protein). The reaction mixture was incubated at 85°C for 1 h and analyzed with HPLC. Abbreviations: OP, octanoyl-peptide; AU, arbitrary units.



FIG S12. Standard curve of proteins used in the determination of molecular mass with gel filtration chromatography. (A) and (B) Chromatograms of the standard proteins used for calibration curve plotting. The absorbance was monitored at 280 nm. The standard proteins are aprotinin (Apr, 6.5 kDa), ribonuclease A (R, 13.7 kDa), carbonic anhydrase (CA, 29 kDa), ovalbumin (O, 43 kDa), conalbumin (C, 75 kDa) and aldolase (Ald, 158 kDa). (C) Calibration curve. The distribution coefficient (K_{av}), which is calculated from the measured elution volume, against the logarithmic molecular mass Log(M_w) of each protein was plotted.



FIG S13. Ramachandran plot of *Tk***-LpI-N structure model.** The Ramachandran plot obtained for the *Tk*-LpI-N structure model quality assessment showed that 95.9% of the residues was in the most favored regions (red area), 4.1% was under the allowed regions (yellow and light-yellow areas), suggesting a good quality of the structure model. The model of *Tk*-LpI-N was validated by PROCHECK module of UCLA-DOE LAB-SAVES v6.0.

Reference:

1. Nesbitt NM, Baleanu-Gogonea C, Cicchillo RM, Goodson K, Iwig DF, Broadwater JA, Haas JA, Fox BG, Booker SJ. 2005. Expression, purification, and physical characterization of *Escherichia coli* lipoyl (octanoyl) transferase. Protein Expr Purif 39:269-282.