Supplementary Information

Crystal structure of LepI, a multifunctional SAM-dependent enzyme which catalyzes pericyclic reactions

Zhenying Chang[1,2,#], Tamar Ansbacher[3,4,#], Lilan Zhang[1,#], Yong Yang[2,5], Tzu-Ping Ko[6], Guimin Zhang[2], Weidong Liu[1], Jian-Wen Huang[1,2], Longhai Dai[2], Rey-Ting Guo[1,2], Dan Thomas Major[3]*, Chun-Chi Chen[2]*

[1]Industrial Enzymes National Engineering Laboratory, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China.

[2]State Key Laboratory of Biocatalysis and Enzyme Engineering, Hubei Collaborative Innovation Center for Green Transformation of Bio-Resources, Hubei Key Laboratory of Industrial Biotechnology, College of Life Sciences, Hubei University, Wuhan, 434200, China

[3]Department of Chemistry, Bar-Ilan University, Ramat-Gan 52900, Israel

[4]Hadassah Academic College, 7 Hanevi'im Street, Jerusalem 9101001, Israel

[5]Tobacco Research Institute of Hubei Province, Wuhan, 430030, China

[6]Institute of Biological Chemistry, Academia Sinica, Taiwan

[#]These authors contribute equally.

*Corresponding authors: majort@biu.ac.il (DTM); ccckate0722@gmail.com (CCC)

Experimental

Protein expression, purification and crystallization

The gene for *A. flavus* LepI (GenBank, XP_002380253.1) was chemically synthesized and cloned into the pET46 Ek/LIC vector (GENE ray Biotech Co., Shanghai, China). The recombinant plasmids were transformed to *E. coli* BL21trxB (DE3). A single transformant of LepI was grown overnight at 37 °C in LB containing 100 μg/mL ampicillin and 50 μg/mL kanamycin. Six liters of fresh LB medium with ampicillin (100 μg/mL) and kanamycin (50 μg/mL) were inoculated with 60 mL overnight cultures and grown to an OD600 of 0.6. The protein expression was induced by 0.6 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 16 °C for 20 h at 16 °C. The cells were harvested by centrifugation at 6,000 x g and re-suspended in a lysis buffer containing 50 mM Tris, pH 8.0, 500 mM NaCl and 20 mM imidazole. The cells were disrupted by French press (GuangZhou JuNeng Biology and Technology Co. Ltd, Guangzhou, China) and then centrifuged at 16,000 x g to remove cell debris. The cell-free supernatant was loaded onto a lysis buffer-equilibrated Ni-NTA column using FPLC system (GE Healthcare). The column was washed by lysis buffer and the target protein was eluted by using a 20-500 mM imidazole gradient. The LepI-containing fractions were combined and further purified using a DEAE column with a buffer of 50 mM Tris, pH 8.0 containing a gradient of 0-500 mM NaCl. The target protein was eluted at ~200 mM NaCl,

which was eventually concentrated to 10 mg/mL in a buffer containing 50 mM Tris, pH 8.0, 200 mM NaCl. The purity was checked by SDS–PAGE analysis (> 95 % purity).

The expression of selenomethionine-substituted (SeMet) LepI was conducted as described in a previous report,[1] and the protein purification procedures were conducted as used for the native enzyme. The SeMet protein was concentrated to 10 mg/mL in 50 mM Tris (pH 8.0), 200 mM NaCl and stored at -80 °C.

Crystallization and data collection

Initial crystallization screening was performed using 768 different reservoir compositions (Hampton Research, Laguna Niguel, California, USA). All of the crystallization experiments were conducted at 25 °C, using the sitting-drop vapor-diffusion method. In general, 2 μL protein containing 10 mM DTT was mixed with 2 μL reservoir solution in 24-well Cryschem plates (Hampton Research) and equilibrated against 500 μL reservoir solution. Initial crystals were observed within 3 days using Structure Screen Ⅰ 7 (200 mM ammonium acetate, 100 mM sodium citrate pH 5.6, 30 % w/v PEG 4000). The crystallization conditions were then optimized to 100 mM sodium citrate, pH 5.6, 100-600 mM NaCl, 20-30 % PEG 4000. The same crystallization condition was used to grow SeMet-LepI crystals.

Structural determination and refinement

The X-ray diffraction datasets of LepI and SeMet-LepI were collected at beam line BL15A1 of the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan). Crystals were soaked in a cryo-protectant containing 50 mM Tris, 200 mM sodium citrate, 30 % PEG 4000, and 500 mM NaCl prior to data collection. The single wavelength anomalous dispersion (SAD) data set for SeMet- LepI was collected at the peak wavelength of 0.979 Å. All data were processed by using HKL2000 program.[2] Prior to structure refinement, 5 % randomly selected reflections were set aside for calculating R_{free} as a monitor of model quality. Initial phase for SeMet datasets was obtained by using SHELXC/D/E[3] from the CCP4i program suit.[4] The figure of merit (FOM) values were over 0.9 and a better model with most side chains was built by ARP/wARP.[5] The model and map were further improved by refinement using Refmac5[6] and Coot.[7] All figures were prepared using the PyMOL program (http://pymol.sourceforge.net/). The data collection statistics are shown in Table S1.

Docking and reaction modeling studies

The ligand Leporin C (**4**) was docked into the active site of LepI (with bound SAM) using the Glide docking program (Schrodinger, Inc.).[8, 9] Following docking, the LepI:SAM:ligand (**4**) complex was solvated with a 20 Å water sphere centered around the active site, using the CHARMM program and stochastic boundary conditions.[10, 11] Subsequently, the system was

minimized and then slowly heated to 298 K during the course of 25 ps using molecular dynamics (MD). To generate **2** and **3** from **4** in the LepI active site, we employed harmonic biasing potentials along pre-selected reaction coordinates in conjunction with MD simulations. A combined quantum mechanics-molecular mechanics (QM/MM) potential was used,[12] where the QM region contained the ligand (i.e. **1**, **2**, **3**, or **4**). The semi-empirical AM1 Hamiltonian was employed [13], i.e. QM(AM1)/MM. To generate **3**, we defined the reaction coordinates as the difference between two bond distances: R(C6-C15) (formation) and R(O_4-C_{13}) (elongation) (Scheme 1). This simultaneous bond breaking/formation ensures that the product, **3**, is obtained with the correct stereochemistry. Quinone methide, (*E*)-**2**, was created similarly, with bonds C7-C12 and O4-C13 both elongated. Subsequently, the LepI:SAM:ligand (**2** or **3**) system was further equilibrated for 1 ns using MD with the QM(AM1)/MM potential. The equilibration was followed by minimization (200 steps), 1 ps of MD and another minimization (200 steps) using a QM(DFT)/MM potential. The specific flavor of DFT was the M06-2X functional in conjunction with the 6-31+G* basis set.[14, 15] Lastly, **1** was created manually using the minimized coordinates of structure **2**. A proton was transferred from H133 to C5 keto group (H133 serves as a base to deprotonate **1**, see Results and discussion). An additional OH group was added to C7 (Scheme 1). The LepI:SAM complex with bound **1** was minimized as well using QM(DFT)/MM. According to the current docked model, the most likely stereoisomer of **1** is (*R*)-C7 due to its proximity to Asp296.

Reference

1. L. Zhang, C. C. Chen, T. P. Ko, J. W. Huang, Y. Zheng, W. Liu, I. Wang, S. R. Malwal, X. Feng, K. Wang, C. H. Huang, S. T. Hsu, A. H. Wang, E. Oldfield and R. T. Guo, *Angew. Chem., Int. Ed. Engl.*, 2016, **55**, 4716-4720.
2. Z. Otwinowski and W. Minor, *Methods Enzymol.*, 1997, **276**, 307-326.
3. G. M. Sheldrick, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2010, **66**, 479-485.
4. E. Potterton, P. Briggs, M. Turkenburg and E. Dodson, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2003, **59**, 1131-1137.
5. S. X. Cohen, M. Ben Jelloul, F. Long, A. Vagin, P. Knipscheer, J. Lebbink, T. K. Sixma, V. S. Lamzin, G. N. Murshudov and A. Perrakis, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2008, **64**, 49-60.
6. G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2011, **67**, 355-367.
7. P. Emsley and K. Cowtan, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2004, **60**, 2126-2132.
8. R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739-1749.
9. T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, *J. Med. Chem.*, 2004, **47**, 1750-1759.
10. B. R. Brooks, C. L. Brooks, 3rd, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comput. Chem*, 2009, **30**, 1545-1614.
11. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem*, 1983, **4**, 187-217.
12. A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227-249.
13. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902-3909.
14. W. J. R. L. Hehre, P. v. R. Schleyer and J. A. Pople, *John Wiley & Sons: New York: 1986*.
15. Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215-241.

Table

Table S1. Data collection and refinement statistics for LepI crystals

	Se-Met	SAM-bound LepI
Data collection[a]		
Space group	C2	C2
Unit-cell a, b, c (Å)	161.3, 62.8, 113.6	160.4, 62.2, 113.2
α, β, γ (°)	90, 113.2, 90	90, 113.3, 90
Resolution (Å)	25 – 2.40 (2.49 –)	25 – 1.85 (1.95 –)
Unique reflections	41186 (4114)	85513 (8409)
Redundancy	7.1 (6.9)	3.3 (3.2)
Completeness (%)	98.7 (99.8)	97.4 (96.3)
Average I/σ (I)	29.1 (4.4)	19.9 (1.4)
CC1/2[b]	0.982 (0.938)	0.903 (0.689)
Refinement[c]		
Resolution (Å)		25 – 1.85 (1.89 –)
No. of reflections		85349 (6234)
R_{work} (%)		18.7 (32.6)
R_{free} (%)		23.1 (36.9)
Ramachandran plot		
Bond length rmsd (Å)		0.008
Bond angle rmsd (°)		1.34
Ramachandran plot		
Favored (%)		98.1
Allowed (%)		1.9
Outlier (%)		0
Average B ($Å^2$)/atoms		
Protein		32.7 / 6096
Ligand		28.6 / 54
Water		41.2 / 779
PDB ID		5ZZD

[a]Numbers in parentheses are for the outermost resolution shell.

[b]CC(1/2) = percentage of correlation between intensities from random half-datasets. Karplus & Diederichs (2012), Science 336, 1030-33.

[c]All positive reflections were used in the refinement except for 5% randomly selected reflections set aside to calculate R_{free}.

Figure

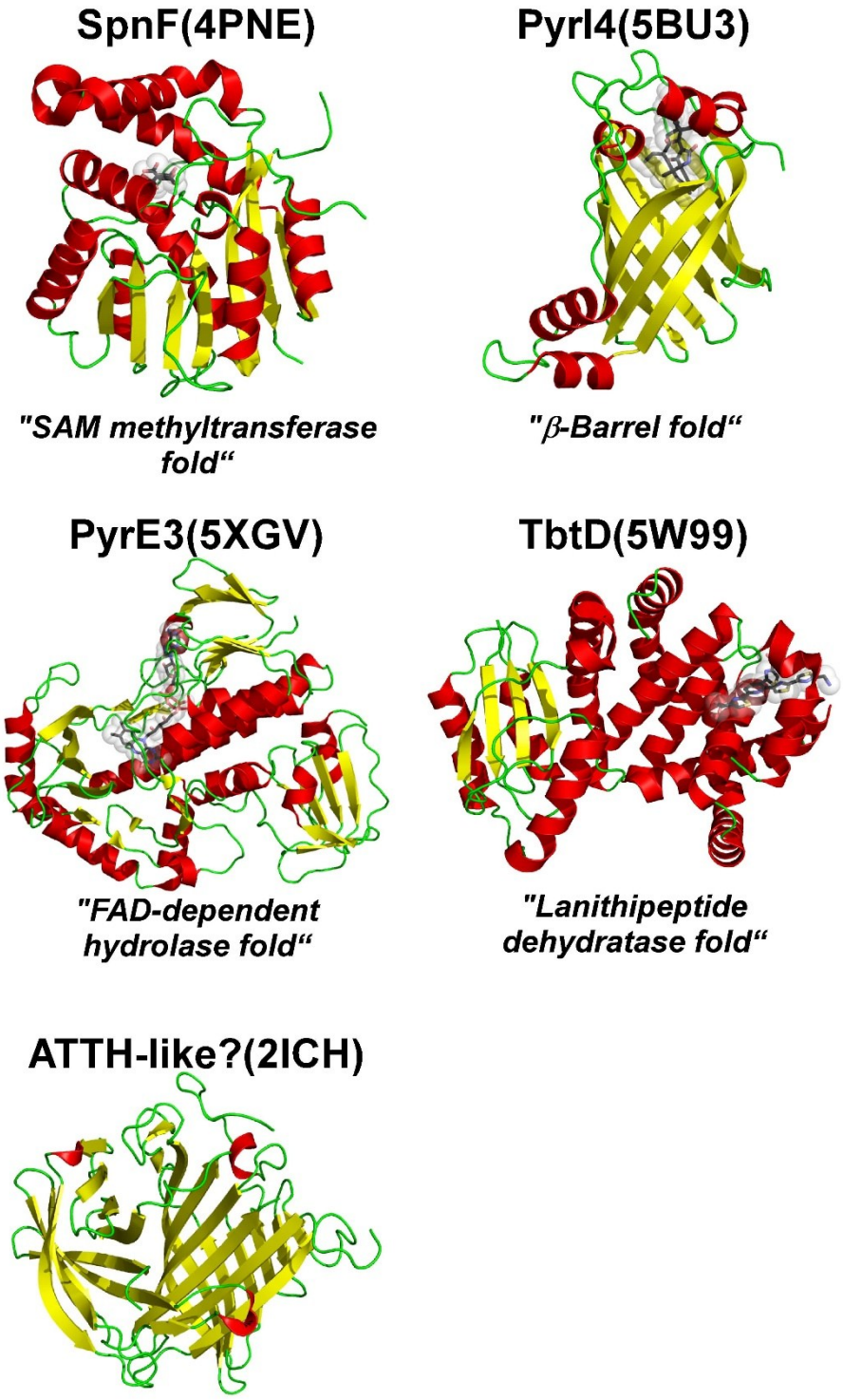

Fig. S1 Structure of various DAases. Cartoon diagrams of five subgroups of DAases are displayed. The bound ligands are shown in sticks. PDB IDs of each structure are labeled in the parentheses.

No:	Chain	Z	rmsd	lali	nres	%id	Description
1:	5w7r-A	36.0	3.1	375	397	23	OXAC;
2:	5w7p-A	35.9	3.2	378	397	22	OXAC;
3:	5w7s-A	35.9	3.2	378	395	22	OXAC;
4:	3gwz-C	32.1	2.6	323	340	20	MMCR;
5:	3gwz-A	32.0	2.5	323	340	20	MMCR;
6:	3gxo-B	31.9	2.6	323	340	21	MMCR;
7:	3gxo-A	31.8	2.6	323	340	20	MMCR;
8:	6c5b-B	31.7	2.3	322	333	23	METHYLTRANSFERASE;
9:	3gxo-C	31.7	2.6	323	340	20	MMCR;
10:	3gwz-B	31.7	2.5	323	340	20	MMCR;
11:	6c5b-A	31.6	2.4	322	333	23	METHYLTRANSFERASE;
12:	3gxo-D	31.4	3.0	323	340	20	MMCR;
13:	3gwz-D	31.3	3.1	323	340	21	MMCR;
14:	4a6d-A	29.6	2.9	327	346	17	HYDROXYINDOLE O-METHYLTRANSFERASE;
15:	5eeh-B	29.4	3.2	322	341	21	CARMINOMYCIN 4-O-METHYLTRANSFERASE DNRK;
16:	4a6e-A	29.4	2.9	327	346	16	HYDROXYINDOLE O-METHYLTRANSFERASE;
17:	5jr3-B	29.3	3.2	322	341	21	CARMINOMYCIN 4-O-METHYLTRANSFERASE DNRK;
18:	5eeh-A	29.3	3.2	322	343	21	CARMINOMYCIN 4-O-METHYLTRANSFERASE DNRK;
19:	1tw3-B	29.2	3.1	322	341	21	CARMINOMYCIN 4-O-METHYLTRANSFERASE;
20:	4wxh-A	29.2	3.1	323	345	21	CARMINOMYCIN 4-O-METHYLTRANSFERASE DNRK;
21:	1tw3-A	29.1	2.8	321	341	21	CARMINOMYCIN 4-O-METHYLTRANSFERASE;
22:	4wxh-B	29.1	3.2	322	341	21	CARMINOMYCIN 4-O-METHYLTRANSFERASE DNRK;
23:	1kyw-F	29.0	3.2	329	362	22	CAFFEIC ACID 3-O-METHYLTRANSFERASE;
24:	3p9k-D	28.9	3.4	318	352	20	CAFFEIC ACID O-METHYLTRANSFERASE;
25:	1tw2-A	28.9	2.8	318	339	21	CARMINOMYCIN 4-O-METHYLTRANSFERASE;

Fig. S2 Partial results of homologous structure of LepI search in Dali database.

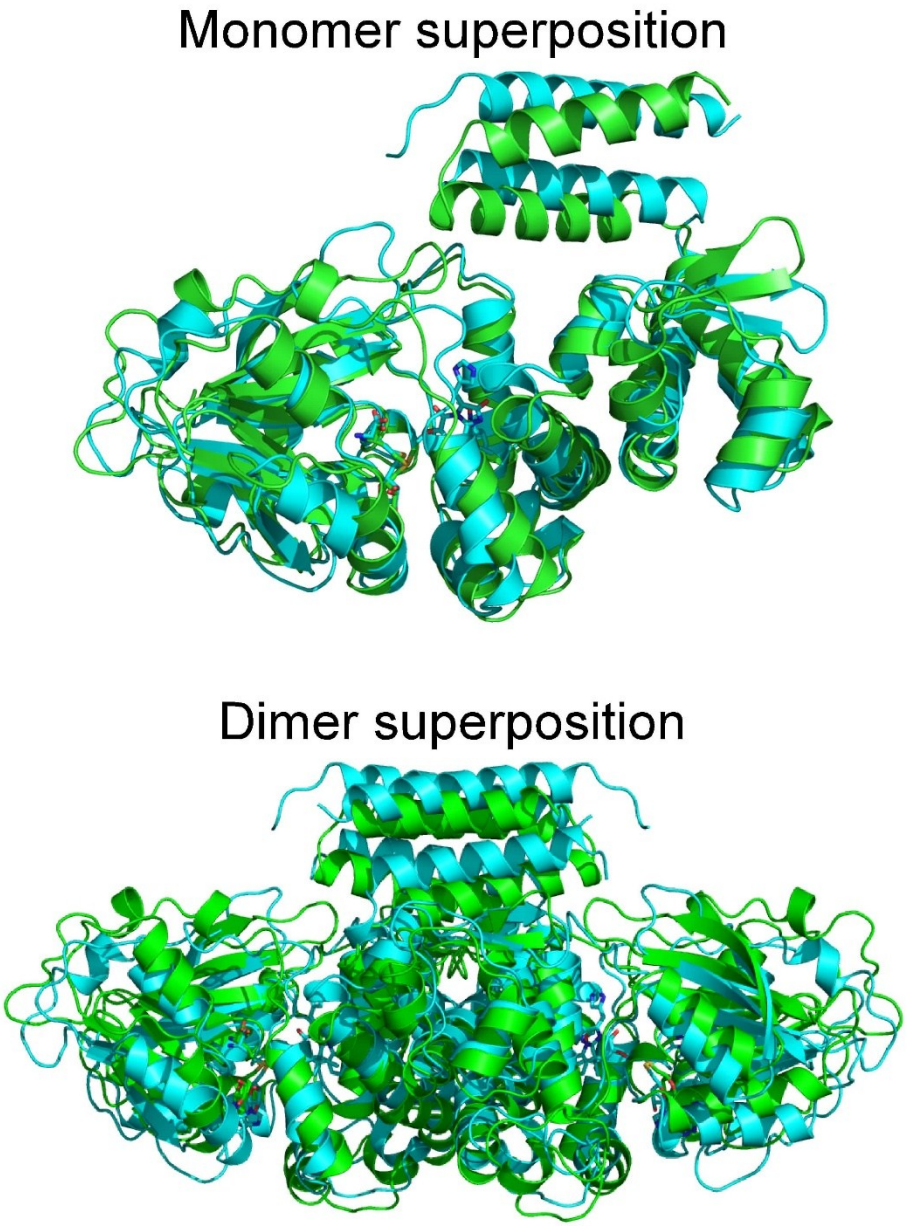

Fig. S3. Superposition of LepI and OxaC. Superposition of (upper panel) monomeric and (lower panel) dimeric configurations of LepI (green) and OxaC (cyan) are shown. The bound ligands are shown in stick models.

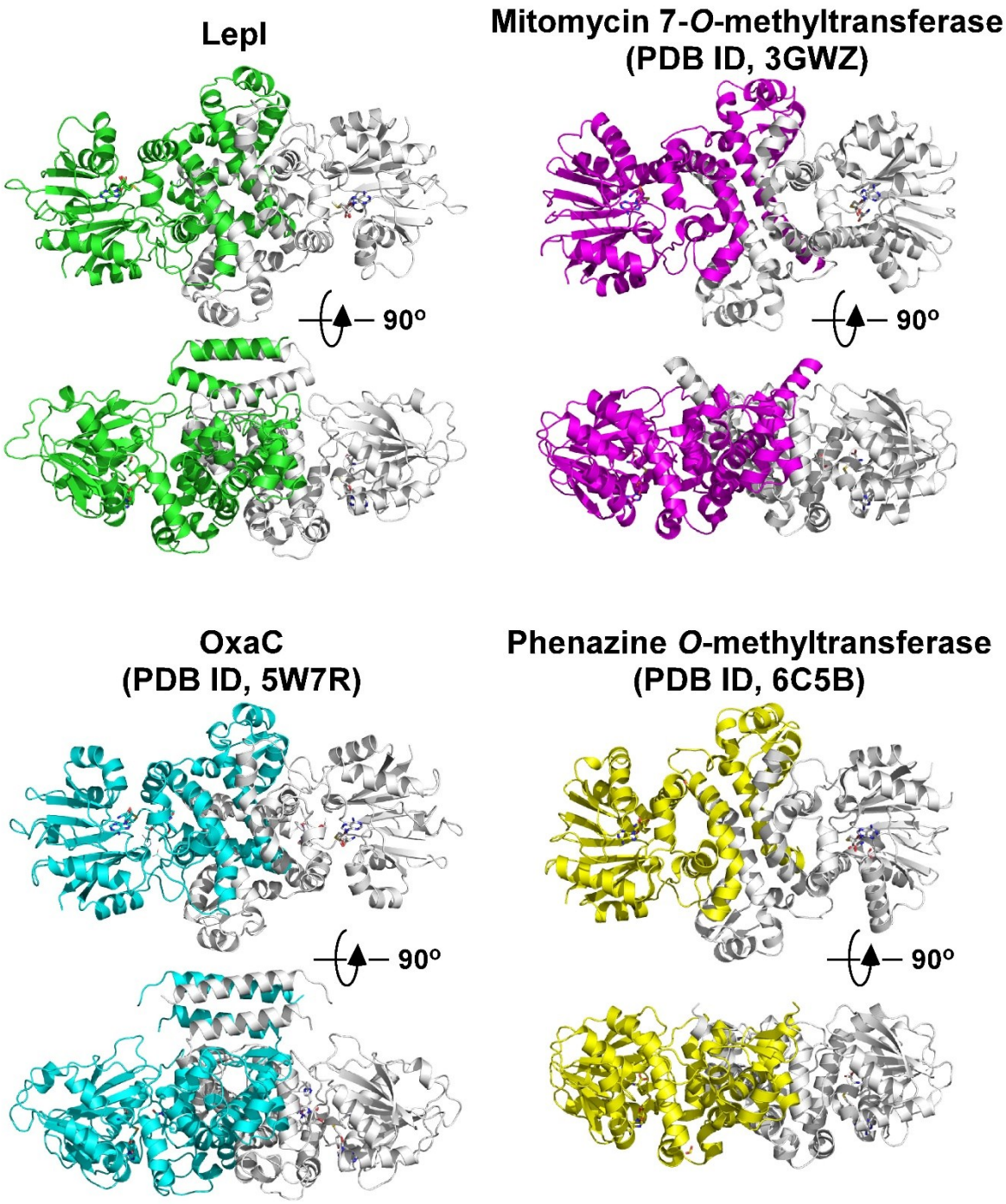

Fig. S4. Dimeric organization of LepI homologous structures. Cartoon diagrams of LepI (color scheme, see Figure 1), OxaC (color scheme, see Figure 1), mitomycin 7-*O*-methyltransferase (two identical polypeptide chains are in magenta and white color) and phenazine *O*-methyltransferase (two identical polypeptide chains are in yellow and white color) are presented. The bound ligands are shown in stick models. Upper and lower images in each panel are related by 90° rotation.

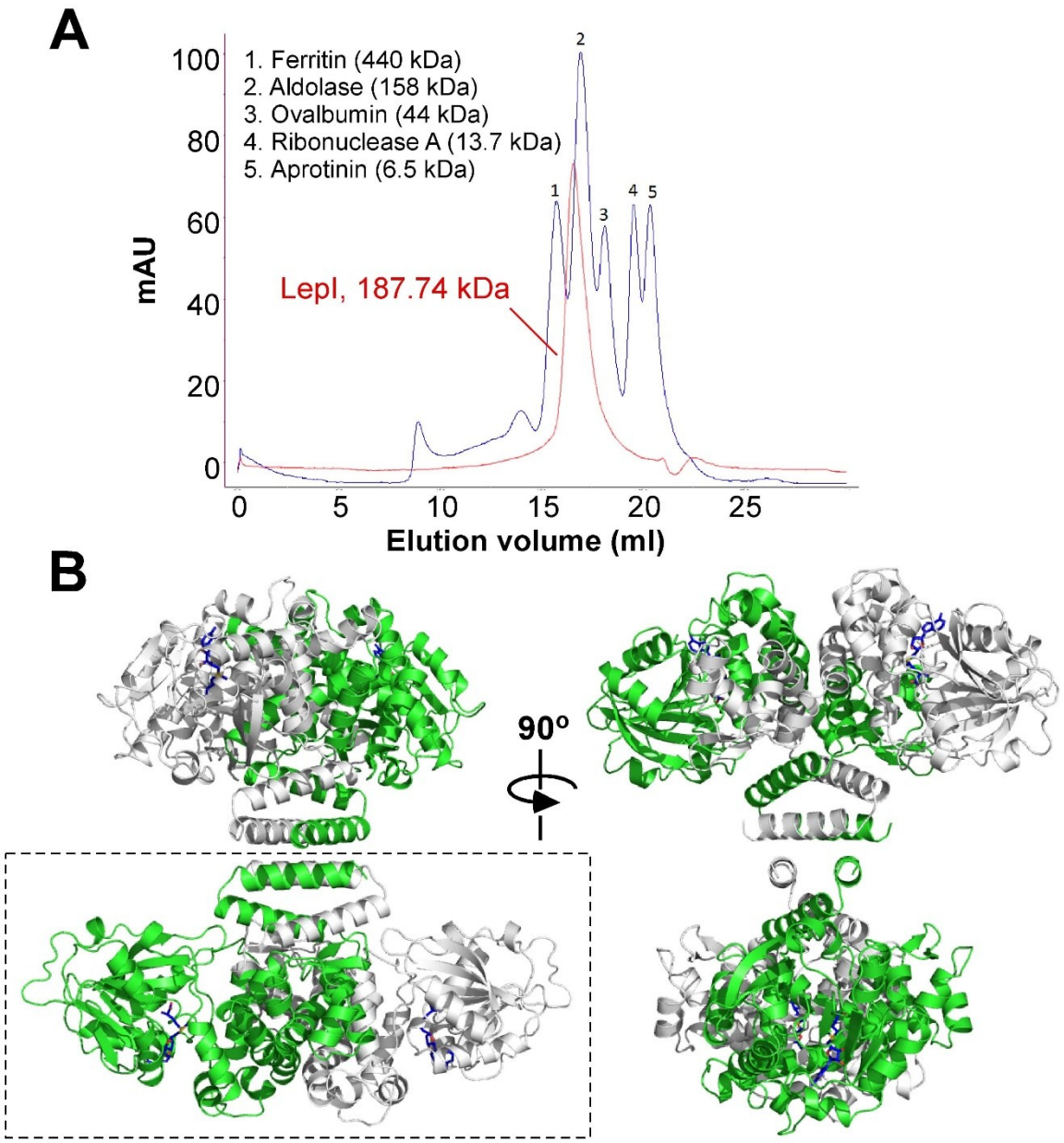

Fig. S5 LepI tetramerization. (**A**) Size exclusion chromatography elution profiles of LepI (red curve), along with the molecular mass standards of 440, 158, 44, 13.7, and 6.5 kDa (blue curve). (**B**) Two LepI dyads that are organized into a tetramer are shown, and two panels are related by 90° rotation. The LepI dimer which resembles other homologous OMT structures are displayed as shown in Fig. 1A (in the dashed line box).

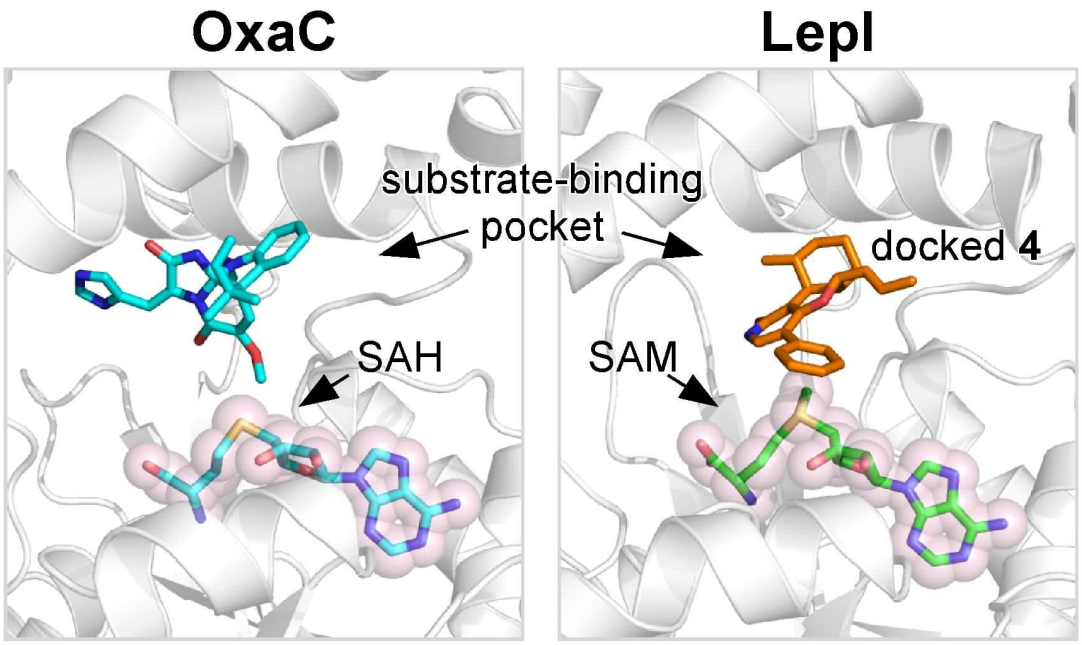

Fig. S6. Substrate binding pocket of OxaC and LepI. Cartoon diagrams of OxaC (PDB ID, 5W7R) and LepI are displayed. Bound ligands of the structures are shown in stick models. The SAH- and SAM-binding site of both structures, substrate-binding pocket of OxaC and the corresponding region of LepI are indicated.

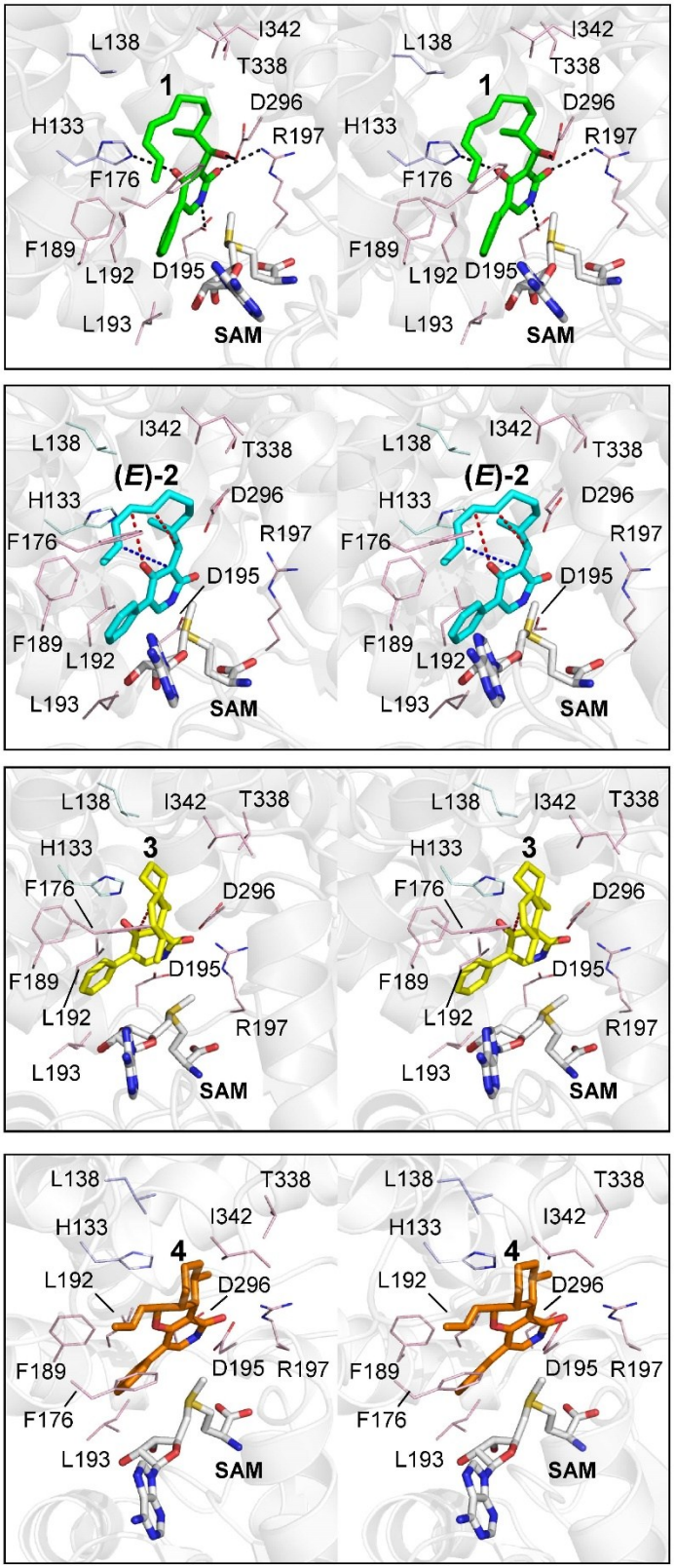

Fig. S7. Stereoview of substrate-interaction networks in LepI. Structural elements, color scheme, dashed line and labels are shown as described in **Fig. 3**.