Structural Basis for Stereoselective Dehydration and Hydrogen-Bonding Catalysis by the SAM-Dependent Pericyclase LepI

Yujuan Cai,[1†] Yang Hai,[2†] Masao Ohashi,[2†] Cooper S. Jamieson,[2] Marc Garcia-Borras,[2] K. N. Houk,[2]* Jiahai Zhou,[1]* Yi Tang[2]*

[1]State Key Laboratory of Bio-organic and Natural Products Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, China; [2]Department of Chemistry and Biochemistry, Department of Chemical and Biomolecular Engineering, University of California Los Angeles, USA.

[†]These authors contributed equally to this work

*Correspondence: yitang@g.ucla.edu, jiahai@mail.sioc.ac.cn, houk@chem.ucla.edu

ABSTRACT

LepI is an *S*-adenosylmethionine (SAM)-dependent pericyclase that catalyzes the formation of 2-pyridone natural product leporin C. Biochemical characterization showed LepI can catalyze the stereoselective dehydration to yield a reactive (*E*)-quinone methide that can undergo bifurcating intramolecular Diels-Alder (IMDA) and hetero-Diels-Alder (HDA) cyclizations from an ambimodal transition state, as well as a [3,3]-retro-Claisen rearrangement to recycle the IMDA product into leporin C. Here we solved the X-ray crystal structures of SAM-bound LepI and in complex with a substrate analog, the product leporin C, and a retro-Claisen reaction transition-state analog to understand the structural basis for the multitude of reactions. Structural and mutational analysis revealed how Nature evolves a classic methyltransferase active site into one that can serve as a dehydratase and a multifunctional pericyclase. Catalysis of both sets of reactions employs His133 and Arg295, two active site residues that are not found in canonical methyltransferases. An alternative role of SAM, which is not found to be in direct contact with the substrate, is also proposed.

Pericyclic reactions are among the most powerful synthetic reactions for making multiple carbon-carbon and carbon-heteroatom bonds in a regio- and stereoselective manner[1]. Despite their prevalence in organic synthesis, only a handful of naturally-occurring enzymes have been characterized to catalyze pericyclic reactions and related [4+2] cycloadditions[2]. Structural characterizations of a few of these enzymes, including chorismate mutase (CM), isochorismate lyase, precorrin-8x methyl mutase, SpnF, etc., showed Nature has evolved a variety of protein folds divergently to accelerate pericyclic reactions and control their regio- and stereoselectivity[3-11].

Recently we discovered a multifunctional *S*-adenosyl-L-methionine (SAM)-dependent *O*-

methyltransferase (OMT)-like pericyclase, LepI, that catalyzes a cascade of reactions starting from the 2-pyridone alcohol **2** to form the dihydropyran-containing fungal natural product leporin C (**10**) (**Figure 1**)[12]. In the absence of LepI, the alcohol **2,** which is derived from the ketoreduction of **1**, can undergo dehydration to give either the (*E*)- or (*Z*)-quinone methide **3** or **4,** respectively. The reactive **3** and **4** can undergo intramolecular Diels-Alder (IMDA) cycloaddition to yield the *endo* products **9** and **6**, respectively, as well as the *exo* adducts **8** and **5,** respectively. The inverse-electron demand hetero-Diels Alder (HDA) cycloaddition of **3** and **4**, on the other hand, affords the desired product **10** and the diastereomeric **7**, respectively. Quantum mechanics (QM) calculations revealed that the activation energies for the formation of these six products from either **3** or **4** are comparable, consistent with **10** being only a minor product in the absence of LepI. We showed that in the presence of LepI, **2** is completely converted to **10**[12]. We also demonstrated computationally that the *endo* IMDA and HDA reactions starting from **3** go through an ambimodal transition state (**TS-1**) and that a post-transition-state bifurcation leads preferentially to the IMDA product in the absence of enzyme catalysis. Therefore, LepI must catalyze the diastereoselective dehydration of **2** to form only **3** and facilitate the subsequent *endo* transition state (**TS-1**) from **3** while suppressing the *exo* transition state (**TS-3**). The active site of LepI can also alter the dynamics of the bifurcating potential energy surface from **TS-1** to favor the HDA product **10** as suggested by computation[12]. Furthermore, we demonstrated that LepI can catalyze the conversion of **9** to **10** via a [3,3]-sigmatropic retro-Claisen rearrangement **TS-2** as a means of kinetically recycling the IMDA shunt **9** to product **10**. While our biochemical and computational studies unveiled these unexpected roles of an enzyme with sequence homology to an OMT, the mechanistic details of catalysis and selectivity achieved by LepI were not known.

Here we report a collection of X-ray crystal structures, which, together with computational

studies, uncover the origins of catalysis and stereoselectivity of LepI. We obtained X-ray crystal structures of *holo*-LepI i) bound with SAM and complexed with ii) the dehydration reaction substrate analogue **1**, iii) product leporin C **10**, and iv) the retro-Claisen reaction substrate analogue **8**. Our structures provide insight into the molecular mechanism of LepI catalysis, including the *E*-stereoselectivity of dehydration, the *endo*-stereoselectivity of IMDA/HDA reactions, and the rate enhancement of the retro-Claisen reaction. A possible divergent evolutionary relationship with SAM-dependent OMTs and a new role of SAM in catalyzing pericyclic reactions are also proposed.

RESULTS

Structure of LepI

The structure of LepI was determined by single anomalous diffraction on a Se-Met derivative crystal diffracting to 2.58 Å resolution, and the phase was extended to a native crystal with 2.14 Å resolution by molecular replacement (**Supplementary Table 1**). The overall tertiary structure of LepI consists of an *N*-terminal dimerization domain (residues 1-122) and a *C*-terminal catalytic domain (residues 123-387) that adopts the typical α/β Rossmann fold observed in all class I SAM-dependent methyltransferases (**Figure 2a**)[13]. In the asymmetric unit, two LepI monomers intertwine through the *N*-terminal dimerization domains with a dimer interface of ~5,000 Å^2 surface area per monomer as calculated by PISA (**Figure 2b**)[14]. Such intimate homodimer architecture is similar to that of many OMTs[2]. Unique to LepI dimer is a domain-swapped four-helix bundle at the *N*-terminus. This leucine-rich coiled-coil enables LepI dimers to pack against each other and form a dimer of dimers, burying ~1,800 Å^2 surface area between the two dimers (**Supplementary Figure 1**). In comparison, the oxaline OMT OxaC[15], the closest homologue to LepI as identified by Dali[16], also harbors a similar four-helix bundle at the *N*-

terminus, but it is not domain-swapped and does not further dimerize. The tetramer oligomerization of LepI is consistent with sedimentation velocity experiment (**Supplementary Figure 2**), which showed that LepI exists in an equilibrium between dimer and tetramer in solution.

The cofactor SAM is copurified with LepI with >90% occupancy when it is overexpressed in *E. coli*[12]; no exogenous SAM was supplemented during protein purification and crystallization steps. Electron density unambiguously delineates the binding of SAM in each LepI monomer at the canonical SAM binding site (**Figure 2c**). LepI makes numerous contacts with SAM that are observed in bona fide OMTs: the α-amino and α-carboxylate groups are recognized through direct and water-mediated hydrogen bond networks; the diol group from the ribose ring donates two hydrogen bonds to D252; the adenine ring is sandwiched between L253 and F276 via van der Waals and π···π interactions. Notably, the sulfonium methyl group makes an unconventional CH···O hydrogen bond with D296, mediated by a water molecule (designated as **W**, the distance between SAM sulfonium methyl carbon and **W** is 3.4 Å). Such CH···O hydrogen bonds involving the SAM sulfonium methyl group are prevalent in SAM-dependent methyltransferases and are believed to be important for SAM binding specificity and facilitating the methyl transfer reaction[17]. However, water molecules have rarely been observed as immediate hydrogen bond acceptors for SAM[18]. The water molecule **W** also makes a hydrogen bond with R295, which resides at the center of a large ~500 $Å^3$ cavity. This putative active site cavity is outlined by both the catalytic domain (monomer A) and the dimerization domain from the other monomer (monomer B) within the homodimer unit (**Figure 2d**). Adjacent to the adenine ring of SAM is a wide substrate entrance tunnel, whereas on the backside of the cavity at the domain interface, multiple water-filled tunnels are found to connect the cavity to bulk solvent (**Supplementary Figure 3a**). Such organization implies that the dimerization event is obligatory for LepI function and that each homodimer is a

minimal integral functional unit.

Structure of LepI-substrate analogue ketone (1) complex (LepI-SAM-1)

To understand how LepI recognizes and orients the substrate alcohol **2** and achieves *E*-selective dehydration to yield **3**, we cocrystalized LepI with ketone **1**, which serves as an unreactive substrate analogue. The structure of this pseudo enzyme-substrate (Michaelis) complex was determined to 1.70 Å resolution. Electron density clearly reveals the binding of **1** in the proposed substrate binding site adjacent to SAM, as well as an ethylene glycol molecule serendipitously positioned on top of the pyridone ring (**Figure 3a, Supplementary Figure 4a**). The atomic resolution of the electron density also allows us to determine the absolute stereochemistry at the chiral C8 of **3** (*S* configuration).

The overall LepI-SAM-**1** structure is very similar to the ligand-free structure with an r.m.s.d. of ~ 0.18 Å for 770 C_α atoms. However, local structural changes were triggered upon substrate binding (**Figure 3b, Supplementary Figure 4b**). The 5-phenyl ring of **1** is locked into a small hydrophobic pocket and participates in favorable π···π interactions with H133 and F189, which cause rotation of F189 side chain. The pyridone ring is recognized via hydrogen bond interactions with H133 and D296 through the 4-OH group and amide nitrogen, respectively.

Unexpectedly, instead of being frozen in the near-attack conformation (NAC) for the cycloaddition step immediately following dehydration, the diene alkyl chain of **1** is captured in an extended conformation projecting into a hydrophobic cleft mainly surrounded by residues from monomer B within the homodimer (**Supplementary Figure 5**), and accommodation of this linear alkyl chain caused a ~3 Å shift of M45(B) side chain. Meanwhile, the space left for the diene chain

in the NAC is now occupied by the ethylene glycol molecule. The electron density reveals that the diene moiety adopts an alternate conformation (*s*-trans and *s*-cis). In both conformations, the diene could participate in favorable π···π interactions with F41(B) (Figure 3a).

The C7 carbonyl group is oriented in parallel with the amide carbonyl facing towards the aforementioned solvent-filled channel and forms an extensive water hydrogen bond network mediated by residues T338, R197, R295, and E350 (Supplementary Figure 3b). Since the C7 carbonyl group closely mimics the alcohol of substrate **2** that leaves in the elimination reaction, such orientation dictates that substrate **2** must undergo *anti*-elimination exclusively to yield (*E*)-quinone methide **3**, which explains our previous biochemical characterization that the enzymatic reaction products were derived from (*E*)-quinone methide intermediate.[12] The water hydrogen bond network helps to protonate the leaving alcohol group. The byproduct water may be transiently trapped by the nearby polar residues and eventually exported to bulk solvent through the solvent-releasing channel (**Supplementary Figure 3b**).

Even though the substrate binding pocket is in proximity to the cofactor SAM, no direct contact was observed between SAM and ketone **1** (the distance between SAM sulfonium methyl carbon and the ketone **1** amide carbonyl oxygen atom is about 6.8 Å). Instead, the water molecule **W** remains to bridge the SAM sulfonium methyl group with D296 and R295. A structure of LepI-SAH-**1** ternary complex (1.84 Å resolution) was also determined during the course of the study, in which *S*-adenosyl-L-homocysteine (SAH), instead of SAM, was identified in the active site (**Supplementary Figure 6**). Even though SAH lacks the sulfonium methyl and a formal positive charge of SAM, the overall structure is essentially identical to the LepI-SAM-**1** complex with an r.m.s.d. of ~ 0.18 Å for 773 C_α atoms., including the active site organization, waters, and substrate

binding mode, except that the diene exclusively adopts the *s*-trans conformation; an extra glycerol molecule was also found at the exit of the solvent-releasing channel. The similarity between these two complexes suggests that the contrasting effects of SAM vs. SAH on LepI catalysis previously observed are not attributed to a structural role of SAM or any allosteric effects but rather to the hydrogen bonding and positive charge of SAM that is absent in SAH.

Structure of LepI enzyme-product (10) complex (LepI-SAM-10)

To gain more insight into how LepI catalyzes cycloaddition, we next determined the crystal structure of LepI-SAM complexed with leporin C (**10**) to 1.78 Å resolution. As expected, no major conformational change was introduced to LepI upon binding of the product. The crystal structure of enzyme-product complex reveals that **10** is accommodated in a manner similar to **1** in the LepI-SAM-**1** complex (**Figure 3c**), through hydrophobic and hydrogen bond interactions. Notably, residue R295 now adopts a different conformation to form a bifurcated hydrogen bond with the amide carbonyl of **10**. The amide carbonyl group concurrently accepts a hydrogen bond donated from a water molecule (**W'**), which likely originates from the leaving alcohol group in the dehydration. This water molecule also participates in the hydrogen bond network at the solvent-releasing channel. The direct hydrogen-bond interaction between the amide carbonyl and guanidinium group of R295 is reminiscent of the electrostatic catalysis predicted previously by our computational study: a sulfonium ion or ammonium ion interacts with the amide carbonyl in the ambimodal **TS-1**[12]. Such electrostatic catalysis was predicted to decrease the C-O bond length and increase the C-C bond length in **TS-1**. This contributes to the observed preference for LepI to catalyze the HDA reaction versus *endo*-mode IMDA in the presence of LepI as compared to nonenzymatic reaction in water. In addition to the amide carbonyl, a lone pair from cyclic ether

oxygen of **10** is shown to accept a hydrogen bond from the side chain of H133. Considering that the general base H133 is inevitably protonated after dehydration in the reaction cascade, H133 is most likely present as the imidazolium form, which makes it an ideal hydrogen-bond donor for catalysis in the following pericyclic reactions. The positively-charged imidazolium ring of H133 maintains hydrogen bonding with the C4-carbonyl of **3** and **TS-1** to enhance the electrophilicity of the pyridone ring. In the [3,3]-sigmatropic retro-Claisen rearrangement, this interaction also stabilizes the developing partial oxyanion in **TS-2**. Together, hydrogen bonds from R295 and H133 to C2 and C4 carbonyl groups activate the quinone methide **3** and the retro-Claisen substrate **9** by stabilizing the polarized transition state.

Superimposing the LepI-SAM-**10** complex onto LepI-SAM-**1** complex reveals that cycloaddition of the diene chain causes the conformational change of M45(B), which swings back to the original position as seen in the unliganded structure (**Figure 3d**). The previously observed ethylene glycol molecule overlays well with C14-C16 of **10**, which illuminates how the linear portion of quinone methide **3** rearranges to the NAC preceding **TS-1**: the diene alkyl chain of quinone methide **3** is induced by the enzyme binding site to rotate and fold on top of the 2-pyridone ring, leaving the C8-C13 region preorganized into a chair-like conformation with C8 methyl group in the equatorial position and the diene moiety in *s*-cis conformation. Accordingly, we docked the ambimodal *endo* **TS-1** and the [3,3]-sigmatropic retro-Claisen rearrangement **TS-2** into the LepI-SAM-**10** structure omitting ligand **10** (**Supplementary Figure 7**). Including the two well-defined water molecules (**W** and **W'**) and SAM in the structure greatly facilitated docking the ligands in productive poses, thereby supporting the hypothesis that they may play important structural roles in defining the active site binding environment.

The basis for LepI diastereoselectivity (*endo* vs *exo*) in the IMDA/HDA reaction was

further supported by modeling the hypothetical LepI-SAM-**TS-3** complex with the assumption that the 5-phenyl group and pyridone ring are bound in a similar fashion as **TS-1** to maintain similar interactions with LepI (**Supplementary Figure 7**). To allow for the *exo* **TS-3** to bind, the hydrophobic diene group (C13-C14) is positioned on the hydrophilic side of the active site, which may potentially cause steric clashes with water **W'** and R295; whereas in the *endo* **TS-1** model, the diene is positioned on the hydrophobic side of the binding site. Thus, the diastereoselectivity likely arises from preference of the diene to avoid the hydrophilic "wall" formed by polar residues (D296, R295, R197, T338) and **W'**. More insights regarding LepI binding of retro-Claisen reaction substrate **9** were provided by the structure of the LepI-SAM-**8** ternary complex (1.66 Å resolution), where **8** is the unreactive spirocyclic product of nonenzymatic exo TS-3 serving as a mimic of the diastereoisomer spirocyclic **9** (see **Supplementary Figure 8** and **Supplementary Text)**.

Catalytic mechanism of LepI

To verify our interpretations of likely modes of catalysis, we performed site-directed mutagenesis on the implicated residues and examined the enzyme activity in *A. nidulans*. In the presence of upstream enzymes that produce the alcohol **2**, a fully active LepI yields only the product **10**, while a completely inactive LepI (or absence of LepI) leads to a mixture of **5-10** (**Figure 4a**). Consistent with our structural conclusion, removal of the essential general base H133 is detrimental to LepI activity; all four mutations (H133A, H133F, H133Q, H133N) completely abolished enzyme activity (**Figure 4a**). As observed in the LepI-SAM-**1** structure, R295 may act as a general acid and indirectly protonate the leaving alcohol. Substitution of R295 with Ala/Gln/Phe/Tyr significantly compromised the dehydration activity, but it was effectively rescued by R295H/K/N. Mutation of other residues structurally implicated in the protonation step during

dehydration (T338A, T338S, R197A, and R197K) did not affect the activity, suggesting that they are not essential for catalysis. Residue D296 plays a major role in binding and polarizing the substrate in order to lock it in the 2-pyridone form rather than its tautomeric 2-hydroxypyridine form. Substitution of residue D296 with Ala led to notable increase in shunt products **5-9**, and the activity can be restored by D296N and D296E, both of which can accept similar hydrogen bonds from the pyridone NH. Although substantial conformational change was observed on M45 accompanying the cycloaddition of **3**, M45A mutation did not affect LepI function *in vivo*, which strongly suggests that folding of the diene chain of **3** to the NAC preceding **TS-1** occurs spontaneously.

To examine the contribution of these active site residues to the catalysis of retro-Claisen rearrangement reaction, we assayed the enzymatic activity of the purified mutants using **9** as the substrate *in vitro* (**Figure 4b**). Consistent with the *in vivo* results, mutation of non-catalytic residues (T338A, M45A) did not affect the pericyclic activity. Alanine substitution for the substrate binding residue D296 caused a 10-fold loss of activity, although the activity was restored to 75% and 90% of the wild-type LepI by replacing A296 with N and E, respectively. In contrast, mutation of the hydrogen bonding catalyst H133 significantly compromised the pericyclase activity of LepI: H133F decreased the activity by 200-fold, and replacing H133 with other neutral hydrogen-bond donors (e.g. H133Q and H133N) did not restore the activity, which demonstrates that the positively-charged imidazolium group is essential for LepI to catalyze the retro-Claisen rearrangement. Substituting R295 with Ala or Gln also completely abolished the activity, while R295N and R295K mutants were active at 10% level of the wile type. Introducing His at position 295 improved the R295A/N mutant LepI activity to 30% of the wild type, suggesting that the partial rescue of inactive mutants by R295H may be attributed to the imidazolium ring that can be

stabilized by the nearby second-shell residue E350 and donate a favorable hydrogen bond to the amide carbonyl to catalyze the retro-Claisen rearrangement. R197, which makes a water-mediated (W') hydrogen bond to the same amide carbonyl, plays a minor catalytic role in this reaction, as R197A mutant retained 50% activity.

Taken together, this structure-activity relationship study provides mechanistic insights into LepI catalysis (**Figure 5**). The dehydration step likely proceeds via the E1-cb mechanism: H133 acts as the key general base to deprotonate the 4-OH group and stabilize the corresponding enolate anion intermediate, while the leaving hydroxyl group is ejected *anti*-periplanar and protonated by water molecules from the water hydrogen bond network, a process that is facilitated by R295. The leaving water molecule (**W'**) could be transiently trapped by these residues and maintain a hydrogen bond with the amide carbonyl. The lack of a dedicated general acid residue to facilitate dehydration presumably is to prevent the newly-formed highly-reactive (*E*)-quinone methide intermediate **3** from being quenched by **W'**, since a good general acid residue could in principle act as a general base in the reverse reaction. Residue D296 is important for substrate binding and favoring the 2-pyridone tautomer. Overall, the dehydration specificity (1,4-*anti*-elimination) is accomplished through positioning the general base and acid in *trans* configurations and trapping the linear alcohol substrate **2** in the corresponding *trans*-conformer.

In the following cycloaddition reaction, spontaneous bond rotation of diene is driven by filling up the void volume over pyridone ring, and better shape-complementarity in the active site when poised for reaction. The diastereoselectivity was achieved by aligning hydrophilic residues at the *exo* side, which in turn favors the *endo* **TS-1** (hydrophobic effect). Residues H133 and R295 act as hydrogen bond donors to lower the transition-state energy barrier and stabilize both **TS-1** and **TS-2**. In particular, the highly polarized transition state **TS-2** (i.e. the oxyanionic 4-carbonyl

oxygen) can be electrostatically stabilized by the positively-charged imidazolium side chain of H133. The cofactor SAM is required for defining the active site shape but not for directly interacting with the substrate. The catalytic role of SAM during the reaction as shown by previous biochemical experiments, and the necessity of a positively-charged analog to effect catalysis is indicative of the electrostatic role of the SAM sulfonium moiety for catalysis, but it is too remote from the substrate for direct hydrogen bonding.

DISCUSSION

A common theme found in pericyclases[2] that catalyze [4+2]-additions is that cyclization follows an initial priming/deprotection reaction that activates the substrate to create/unmask the reactive functional groups[21]. This priming/deprotection reaction may be catalyzed by a separate enzyme (e.g. tandem reaction catalyzed by SpnM and SpnF) or the cyclase itself (e.g. solanapyrone synthase), which generates the reactive intermediate *in situ*. The second scenario is most often encountered when the liberated intermediate is too reactive and readily forms undesired cyclized products without enzymatic control. Therefore, by devising such a cascade scheme, the unstable intermediate generated *in situ* undergoes cyclization immediately inside the cyclase.

LepI also follows this principle, and our studies here provide the structural basis and mechanistic insight into LepI catalysis, such as 1) positioning the general base (H133) and general acid (water facilitated by R295) in the *trans*-configuration to cause the substrate **2** to undergo stereoselective *anti*-elimination to yield (*E*)-quinone methide **3**; 2) setting the "amphiphilic" active site to favor the *endo*-conformation during cycloaddition; 3) repurposing the two cationic residues (H133 and R295) as hydrogen bonding catalysts to lower the transition state barrier and to electrostatically stabilize the polarized transition state; 4) repurposing SAM that normally

functions as a methyl donor in methyltransferases to play an electrostatic role in non-methyltransfer catalysis.

In both natural and artificial enzymes that catalyze Diels-Alder reactions, it has been found that a strategy for rate enhancement is to donate specific hydrogen bonds to the dienophile in order to lower its LUMO energy[22,23] and to stabilize the polarized transition state. In LepI catalysis, H133 and R295 are vivid examples of hydrogen bond catalysts that simultaneously donate hydrogen bonds to both the amide carbonyl and quinone methide carbonyl groups. Such hydrogen bonding effectively lowers the LUMO energy of quinone methide and stabilizes the transition state that involves charge transfer to this moiety.

While many standalone pericyclases utilize binding energy to preorganize the linear substrate in a restricted NAC to overcome the rotational entropic barrier, LepI apparently does not function as an "entropy trap" to increase reaction rate[7,8,10,21,22,24,25]. The undehydrated substrate is bound in the extended conformation by LepI in order to minimize the $\pi\cdots\pi$ repulsion between the diene and the electron-rich phenolate prior to dehydration. The slight loss in entropy that occurs when the linear substrate rotates from extended conformation to the restricted NAC can be compensated by the gain of enthalpy from interaction between the diene and electron-deficient quinone methide and the enthalpic change that eventually results from the cycloaddition reaction. This may be a general feature of all dehydration-promoted pericyclases.

This feature is also reminiscent of another well-characterized pericyclase, chorismate mutase (CM), which catalyzes the [3,3]-sigmatropic Claisen rearrangement of chorismate to prephenate. Unfavorable decreases in $\Delta S^{\ddagger}$ have been consistently observed for natural CMs and catalytic antibodies, which suggests that CM does not act as an entropy trap[26]. Instead, the strategy for CM catalysis is a combination of both reactant destabilization and transition state stabilization[26].

It has become clear that transition state stabilization is the more important contributor to CM catalysis: donation of favorable hydrogen bond from cationic residue (such as Arg) to the vinyl ether oxygen in chorismate can effectively stabilize the developing partial negative charge on transition state[27]. By replacing this critical arginine with its neutral analogue citrulline, which donates hydrogen bond but is incapable of favorable electrostatic interactions, it has been estimated that this positively-charged arginine can contribute 5.9 kcal/mol to transition state stabilization versus 0.6 kcal/mol to the binding energy of the ground state.

Strikingly similar to CM, LepI uses cationic residues (H133 and R295) as major contributor for catalyzing [3,3]-sigmatropic retro-Claisen rearrangement of **9** to **10**. Substrate ground-state destabilization does not contribute, since the spirocyclic substrate **9** is essentially being frozen in the NAC like conformation, and no significant bond rotation is required to preorganize this substrate to be shape complementary to the LepI active site. The major contribution to catalysis must come from hydrogen bonding and electrostatic interaction. Substitution of H133 by neutral hydrogen bonding residue Gln and Asn reduces the catalytic activity by 50-fold. This severe loss of catalytic efficiency implies that H133 exists in the positively charged imidazolium form, which underscores the importance of electrostatic complementarity. The importance of positively-charged R295 to this pericyclic reaction has also been demonstrated by both QM studies and our current mutagenesis study[12]. Consistent with this prediction, R295A/Q mutants are devoid of retro-Claisen rearrangement activity (>1,000-fold decrease, undetectable), whereas activity is partially retained with similar cationic mutants such as R295H (30%) (**Figure 4b**).

Although the precise catalytic role of SAM remains unclear, we have now found that the sulfonium does not contact the amide carbonyl. Nevertheless, the electrostatic effect of the sulfonium or the ammonium analog can stabilize the transition state. Given that the positive charge

of SAM is required for optimal activity of LepI, we propose that SAM serves as an electrostatic catalyst that stabilizes the transition state, which has a higher dipole moment than the reactant. This type of catalysis is well known through the work of Warshel, was demonstrated in solution chemistry by Wilcox, and has gained recent prominence under the name "electric field catalysis" through the work of Boxer, Shaik, Coote, and Head-Gordon[28-33].

Our structural study of LepI also suggests an interesting evolutionary origin, that a methyltransferase ancestor of LepI has been co-opted to a multifunctional dehydratase and pericyclase. The protein folds and SAM binding pockets of LepI are nearly identical to the methyltransferase OxaC (**Supplementary Figure 10**). Comparison of the LepI active site with that of the OMT OxaC, along with structure-based sequence alignment (**Supplementary Figure 11**), reveals how key residues have been evolved to catalyze the reactions shown in Figure 1 instead of a S_N2 methyl transfer reaction. In all functional *O*-MTs, the charge relay system of the His-Glu catalytic dyad is strictly conserved; the histidine acts as the general base that deprotonates the substrate nucleophile (hydroxyl group), and glutamate acts as the second shell residue to elevate the pKa of histidine. In LepI, the glutamate is conserved as E350, whereas the corresponding general base histidine is replaced with arginine. This substitution is beneficial for LepI in two ways. First of all, as discussed above, arginine is a better hydrogen bond donor and electrostatic catalyst in LepI relative to histidine. Secondly, arginine is a poor general base due to its high pKa (12.5) and thus remains protonated in the resting state. Therefore, removing this key general base in LepI may be a strategy to eliminate adventitious methyltransferase activity.

In summary, our structural studies of LepI presented here provide mechanistic insight into enzymatic dehydration-triggered IMDA and HDA reactions, as well as hydrogen bonding and electrostatic catalysis of the retro-Claisen rearrangement. A proper understanding of the molecular

mechanism of LepI is important for designing new pericyclases *de novo* and also provides a template for the discovery of novel functions of OMT-like enzymes.

ACKNOWLEDGEMENT

This work was supported by the NSFC (91856202) and the Strategic Priority Research Program (B) of CAS (XDB20000000) to JZ, NIH (1R01AI141481) and NSF (CHE-1806581) to YT and KNH. Chemical characterization studies were supported by shared instrumentation grants from the NSF (CHE-1048804) and the NIH NCRR (S10RR025631). We thank the staff of beamline BL17U1, BL18U1 and BL19U1 of Shanghai Synchrotron Radiation Facility (China) for access and help with the X-ray data collection. We also thank Prof. Jianhua Gan in Fudan University for help with structure refinement and Prof. Jiafu Long for help with ultracentrifugation sedimentation measurement. The computational resources from the UCLA Institute of Digital Research and Education (IDRE) are gratefully acknowledged. MO is supported by overseas postdoctoral fellowship from The Uehara Memorial Foundation, Japan. YH is a Life Sciences Research Foundation fellow sponsored by the Mark Foundation for Cancer Research. We thank Ms. Eun Bin Go for careful proofreading of the manuscript.

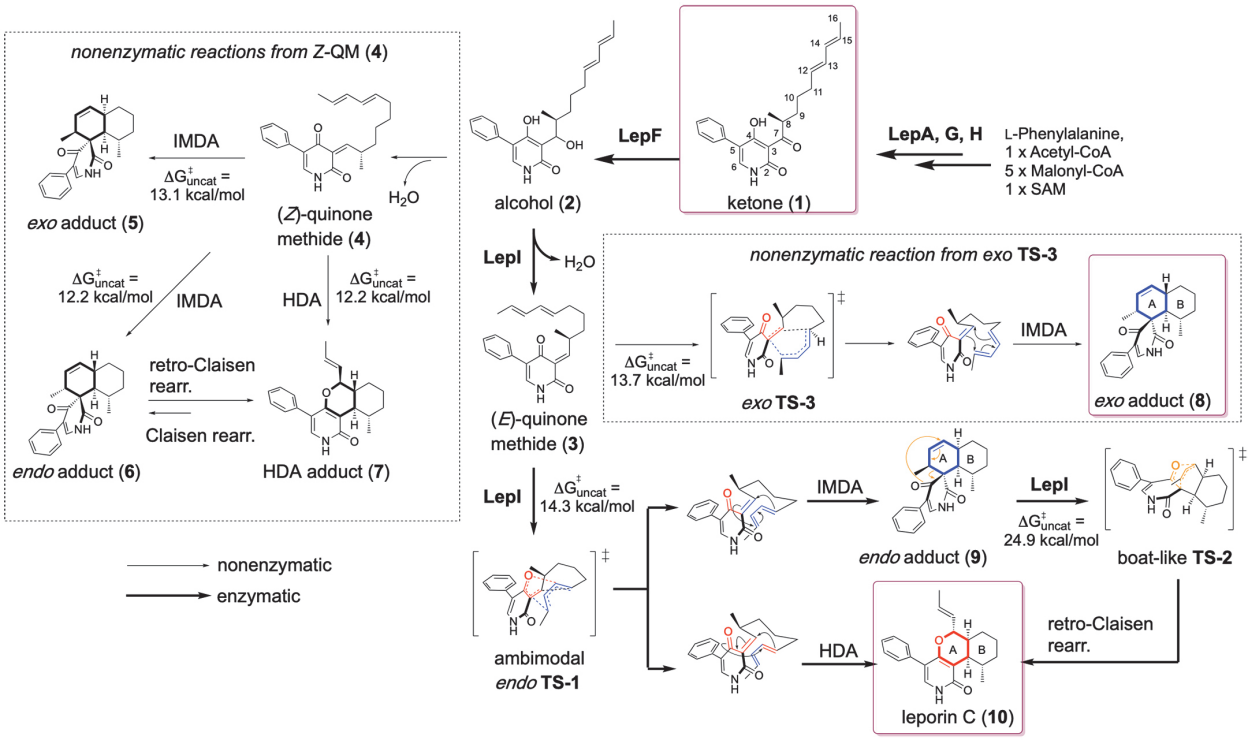

Figure 1 | Leporin C biosynthesis pathway highlighting LepI-catalyzed reaction cascade. In the absence of LepI, spontaneous dehydration of alcohol **2** yields a (*E*/Z)-mixture of quinone methides (**3**/**4**), which nonenzymatically form Diels-Alder and hetero Diels-Alder adducts. The compounds used in structural study are highlighted by maroon boxes.

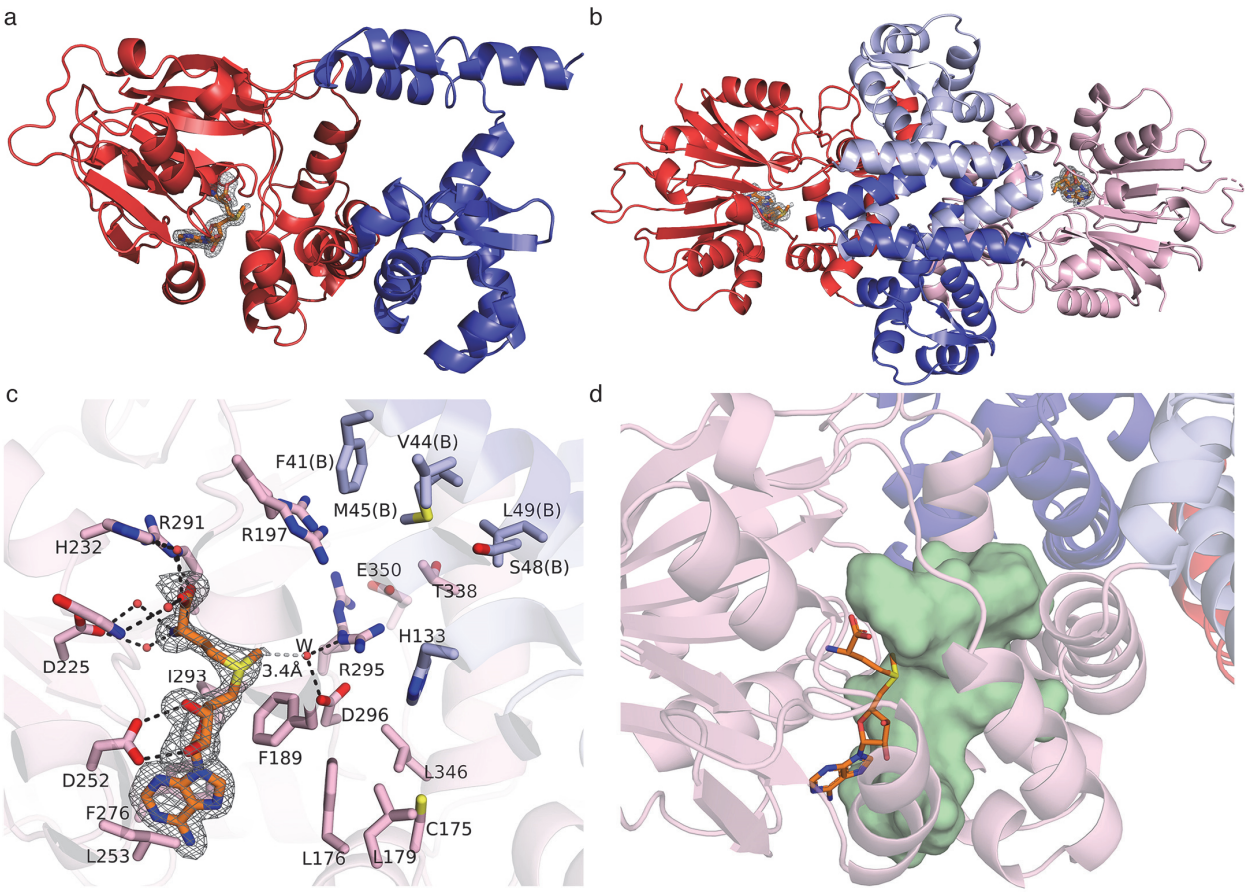

Figure 2 | LepI structure and the SAM binding site. (a) The overall tertiary structure of LepI is shown in cartoon model, with the *N*-terminal dimerization domain and *C*-terminal catalytic domain colored in blue and red, respectively. Simulated-annealing omit map (grey mesh, contoured at 2.5 σ) indicates binding of SAM at the canonical SAM binding site. (b) Intimate LepI homodimer featured by an intertwined dimer interface. (c) Close-up view of SAM binding site. Hydrogen bond interactions are indicated with black dashes. (d) Next to SAM is a large substrate binding cavity and its entrance tunnel (shown together as green surface). The volume of cavity was calculated using POCASA.[19]

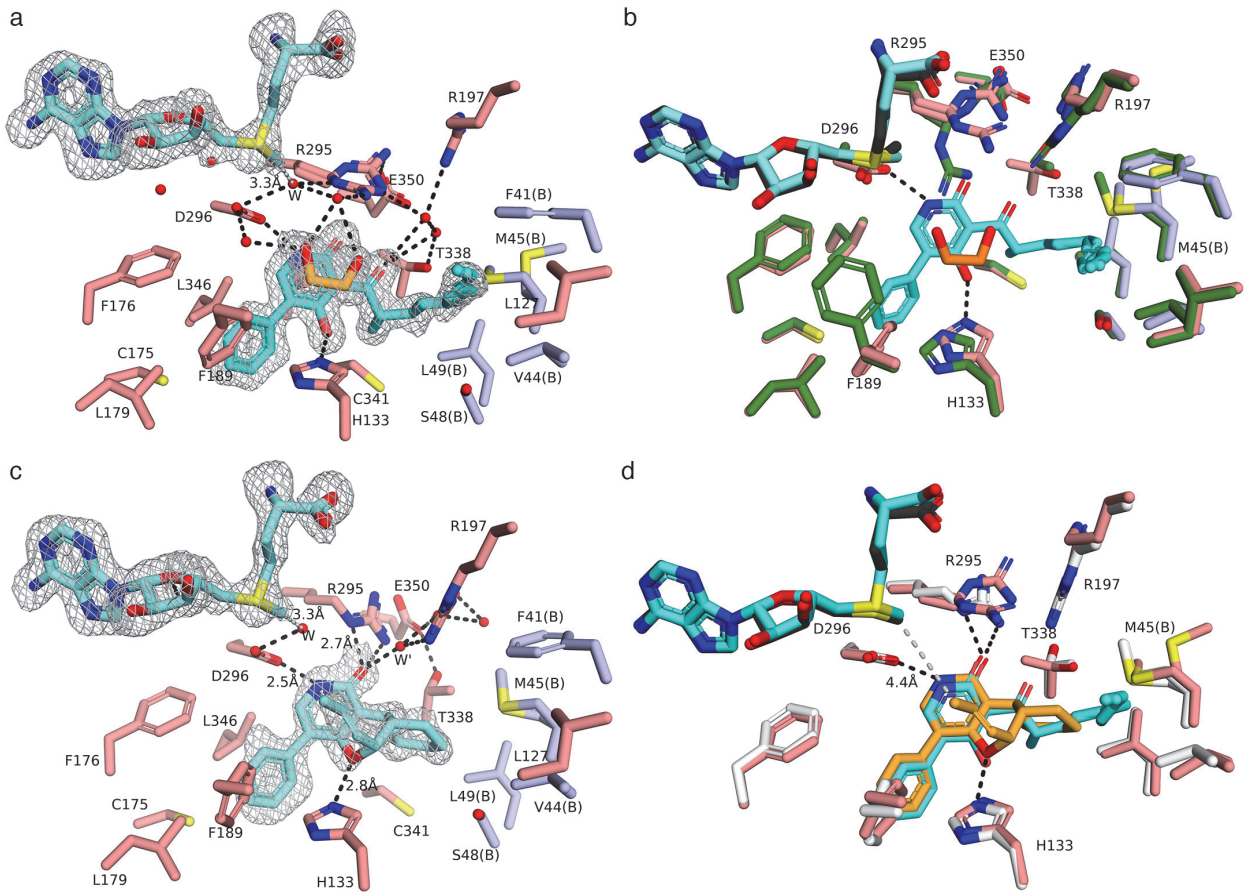

Figure 3 | Crystal structure of LepI pseudo enzyme-substrate complex and enzyme-product complex. Simulated-annealing omit maps are shown in black mesh and contoured at 3.0 σ. Hydrogen bond interactions are indicated with black dashed lines. (a) Crystal structure of LepI-SAM-**1** (pseudo enzyme-substrate complex). Note that two rotamers of R295 and M45(B) are observed, and two conformations of diene (*s*-trans and *s*-cis) are modeled according to the electron density. Residues from monomer A are colored in salmon, whereas residues from monomer B are colored in light blue and labeled with B in parenthesis. (b) Superposition of LepI-SAM-**1** complex (colored as in (a)) with unliganded LepI-SAM (residues are colored in dark green while SAM is colored in black). Substantial conformational changes are observed for F189, R295, R197, and M45(B). (c) Crystal structure of LepI-SAM-**10**. Residues and ligands are color coded as in (a). (d) Superposition of LepI-SAM-**10** (residues are colored in white, SAM is colored in gray, ligand is colored in orange) with LepI-SAM-**1** (color coded as in (a), note that ethylene glycol is not shown here for clarity).

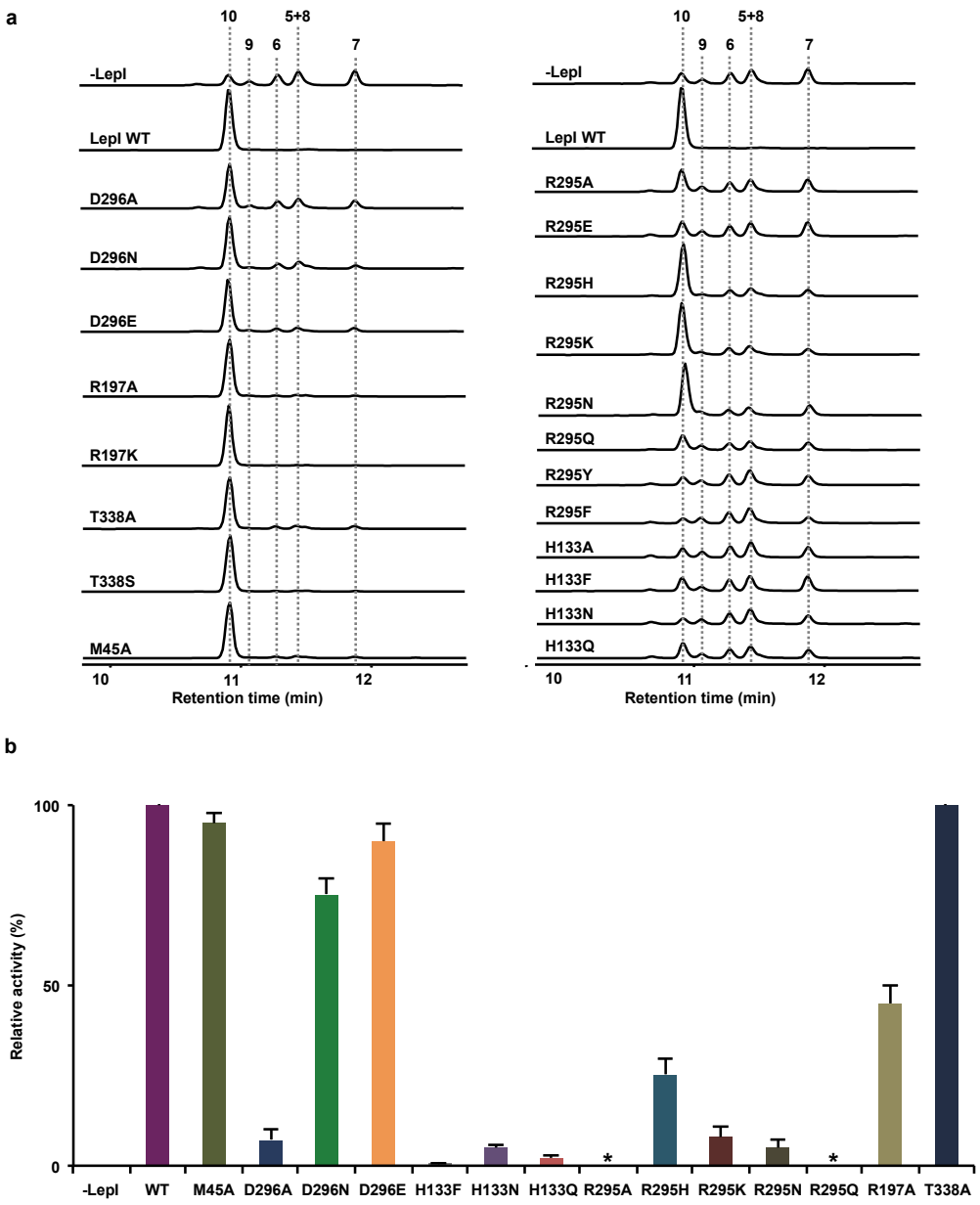

Figure 4 LepI Mutant activity. (a) *In vivo* activity of LepI or mutants to synthesize leporin C (**10**) starting from the alcohol substrate **2**. **2** is synthesized by upstream biosynthetic enzymes; for details of pathway see Ref. 12. (b) *In vitro* retro-Claisen rearrangement activity using **9** as the substrate. Asterisks indicate mutants with no measurable activity. Error bars show s.d. of three independent experiments (n = 3).

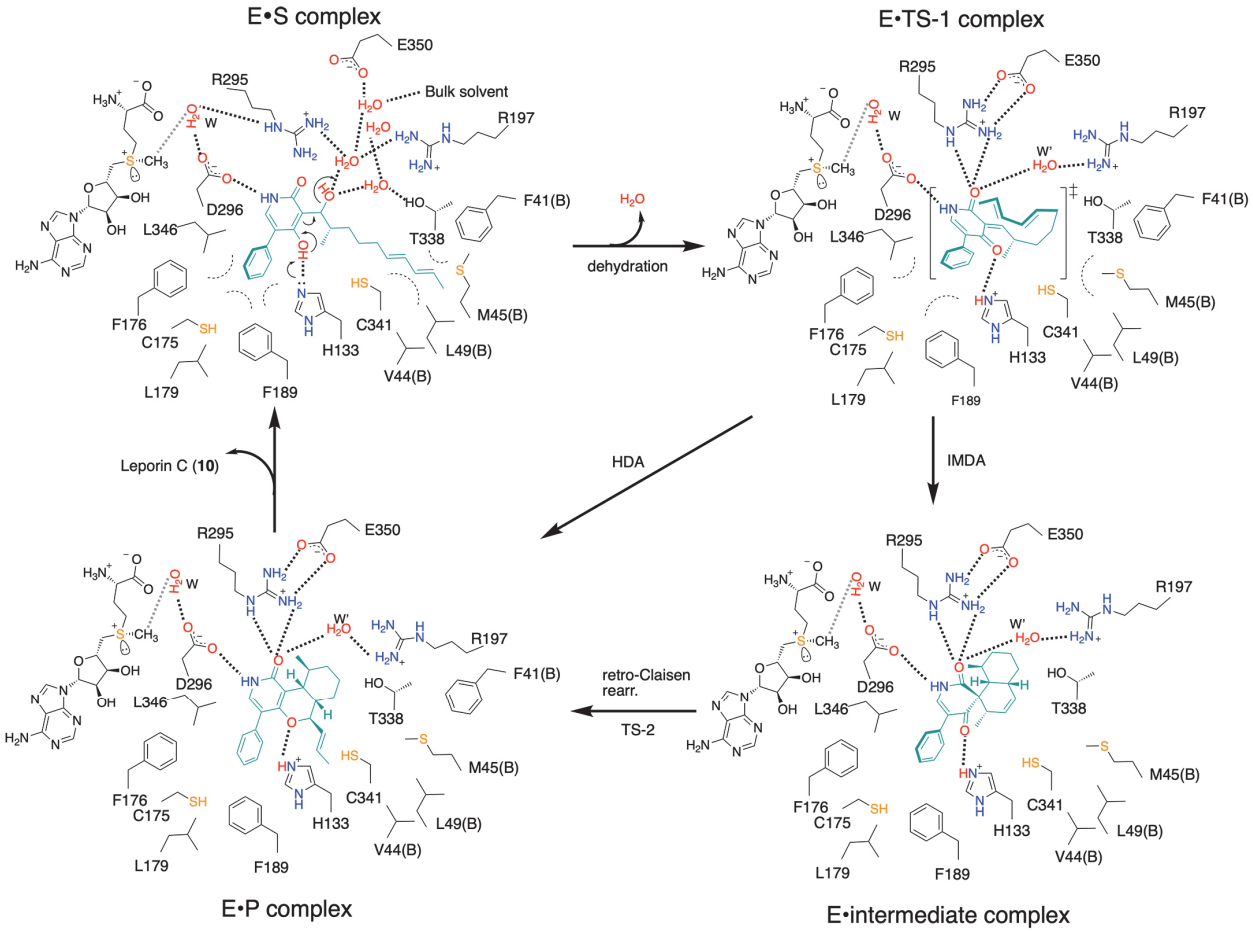

Figure 5 | Proposed catalytic mechanism of LepI. Dashed lines indicate hydrogen bonds.

REFERENCES

1. Nicolaou, K. C., Snyder, S. A., Montagnon, T. & Vassilikogiannakis, G. The Diels–Alder reaction in total synthesis. *Angew. Chem. Int. Ed.* **41**, 1668–1698 (2002).
2. Cooper, S. J., Ohashi, M., Liu, F., Tang, Y., Houk, K. N. The expanding world of biosynthetic pericyclases: cooperation of experiment and theory for discovery. *Nat. Prod. Rep.* Advance Article DOI: 10.1039/c8np00075a (2018).
3. Chook, Y. M., Ke, H. & Lipscomb, W. N. Crystal structures of the monofunctional chorismate mutase from *Bacillus subtilis* and its complex with a transition state analog. *Proc. Natl. Acad. Sci. USA* **90**, 8600–8603 (1993).
4. Fage, C. D. et al. The structure of SpnF, a standalone enzyme that catalyzes [4+2] cycloaddition. *Nat. Chem. Biol.* **11**, 256–258 (2015).
5. Zaitseva, J., Lu, J., Olechoski, K. L. & Lamb, A. L. Two crystal structures of the isochorismate pyruvate lyase from *Pseudomonas aeruginosa*. *J. Biol. Chem.* **281**, 33441–33449 (2006).
6. Shipman, L. W., Li, D., Roessner, C. A., Scott, A. L. & Sacchettini, J. C. Crystal structure of precorrin-8x methyl mutase. *Structure* **9**, 587–596 (2001).
7. Zheng, Q. et al. Enzyme-dependent [4+2] cycloaddition depends on lid-like interaction of the *N*-terminal sequence with the catalytic core in Pyrl4. *Cell Chem. Biol.* **23**, 352–360 (2016).
8. Byrne, M. J. et al. The catalytic mechanism of a natural Diels-Alderase revealed in molecular detail. *J. Am. Chem. Soc.* **138**, 6096–6098 (2016).
9. Cogan, D. P. et al. Structural insights into enzymatic [4+2] azacycloaddition in thiopeptide antibiotic biosynthesis. *Proc. Natl. Acad. Sci. USA* **114**, 12928–12933 (2017).

10. Zheng, Q. et al. Structural insights into a flavin-dependent [4+2] cyclase that catalyzes *trans*-decalin formation in pyrroindomycin biosynthesis. *Cell Chem. Biol.* **25**, 718–727 (2018).
11. Newmister, S. A. et al. Structural basis of the Cope rearrangement and cyclization in hapalindole biogenesis. *Nat. Chem. Biol.* **14**, 345–351 (2018).
12. Ohashi, M. et al. SAM-dependent enzyme-catalysed pericyclic reactions in natural product biosynthesis. *Nature* **549**, 502–506 (2017).
13. Liscombe, D. K., Louie, G. V. & Noel, J. P. Architectures, mechanisms and molecular evolution of natural product methyltransferases. *Nat. Prod. Rep.* **29**, 1238–1250 (2012).
14. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
15. Newmister, S. A. et al. Unveiling sequential late-stage methyltransferase reactions in the meleagrin/oxaline biosynthetic pathway. *Org. Biomol. Chem.* **16**, 6450–6459 (2018).
16. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res*. **38**, w545–549 (2010).
17. Horowitz, S. et al. Conservation and functional importance of carbon-oxygen hydrogen bonding in AdoMet-dependent methyltransferases. *J. Am. Chem. Soc.* **135**, 15536–15548 (2013).
18. Fick, R. J. et al. Water-mediated carbon-oxygen hydrogen bonding facilitates *S*-adenosylmethionine recognition in the reactivation domain of cobalamin-dependent methionine synthase. *Biochemistry* **57**, 3733–3740 (2018).
19. Yu, J., Zhou, Y., Tanaka, I. & Yao, M. Roll: a new algorithm for the detection of protein pockets and cavities with a rolling probe sphere. *Bioinformatics* **26**, 46–52 (2010).

20. Pauling, L. Molecular architecture and biological reactions. *Chem. Eng. News* **24**, 1375–1377 (1946).

21. Jeon, B.-S., Wang, S.-A., Ruszczycky, M. W. & Liu, H.-W. Natural [4+2]-cyclases. *Chem. Rev.* **117**, 5367–5388 (2017).

22. Preiswerk, N. et al. Impact of scaffold rigidity on the design and evolution of an artificial Diels-Alderase. *Proc. Natl. Acad. Sci. USA* **111**, 8013–8018 (2014).

23. Minami, A. & Oikawa, H. Recent advances of Diels-Alderases involved in natural product biosynthesis. *J. Antibiot.* **69**, 500–506 (2016).

24. Breslow, R. Hydrophobic effects on simple organic reactions in water. *Acc. Chem. Res.* **24**, 159–164 (1991).

25. Blokzijl, W., Blandamer, M. J. & Engberts J. B. F. N. Diels-Alder reactions in aqueous solutions. Enforced hydrophobic interactions between diene and dienophile. *J. Am. Chem. Soc.* **113**, 4241–4246 (1991).

26. Lee, A. Y., Stewart, J. D., Clardy, J. & Ganem, B. New insight into the catalytic mechanism of chorismite mutase from structural studies. *Chem. Biol.* **2**, 195–203 (1995).

27. Burschowsky, D. et al. Electrostatic transition state stabilization rather than reactant destabilization provides the chemical basis for efficient chorismate mutase catalysis. *Proc. Natl. Acad. Sci. USA* **111**, 17516–17521 (2014).

28. Warshel, A. et al. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **106**, 3210–3235 (2006).

29. Smith, P. J. & Wilcox, C. S. The chemistry of functional group arrays. Electrostatic catalysis and the “intramolecular salt effect”. *Tetrahedron* **47**, 2617–2628 (1991).

30. Fried, S. D. & Boxer, S. G. Electric Fields and Enzyme Catalysis. *Annu. Rev. Biochem.* **86**, 387–415 (2017).
31. Shaik, S., Mandal, D. & Ramanan, R. Oriented electric fields as future smart reagents in chemistry. *Nat. Chem.* **8**, 1091–1098 (2016).
32. Aragonès, A. C. et al. Electrostatic catalysis of a Diels–Alder reaction. *Nature* **531**, 88–91 (2016).
33. Welborn, V. V., Ruiz, P. L. & Head-Gordon, T. Computational optimization of electric fields for better catalysis design. *Nat. Catal*. **1**, 649–655 (2018).

Author contributions Y.C., Y.H., M.O., K.N.H. J.Z., and Y.T. developed the hypothesis and designed the study. Y.C. and Y.H. purified and crystallized protein samples and solved the X-ray structures. M.O. and Y.H. performed *in vivo* and *in vitro* experiments. M.O. performed compound isolation and characterization. C.S.J, and M.G.-B. performed the computational analysis. All authors analyzed and discussed the results. Y.C, Y.H., K.N.H., J.Z., and Y.T. prepared the manuscript.

Competing financial interests, The authors declare no competing financial interests.

Data availability. Data that support the findings of this study are available within the paper and its Supplementary Information, or are available from the corresponding authors upon reasonable request. The structural factors and coordinates of LepI-SAM complex, LepI-SAM-**1** complex, LepI (C52A)-SAH-**1** complex, LepI (C52A)-SAM-**10** complex, and LepI (C52A)-SAM-**8** complex have been deposited in the Protein Data Bank under the IDs 6IX3, 6IX5, 6IX7, 6IX9, and 6IX8, respectively.