Discovery of enzymatic Alder-ene reaction and origins of catalytic selectivity

Masao Ohashi,[1,6] Cooper S. Jamieson,[2,6] Yujuan Cai,[3,6] Dan Tan,[1] Daiki Kanayama,[2] Man-cheng Tang,[1] Sarah M. Anthony,[2] Jason V. Chari,[2] Joyann S. Barber,[2] Elias Picazo,[2] Thomas B. Kakule,[1] Shugeng Cao,[4,5] Neil K. Garg,[2] Jiahai Zhou,[3]* K. N. Houk[1,2]* & Yi Tang[1,2]*

[1]Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, Los Angeles, California, USA.
[2]Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, California, USA.
[3]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, Shanghai, China.
[4]Department of Pharmaceutical Sciences, Daniel K. Inouye College of Pharmacy, University of Hawaii at Hilo, 200 West Kawili Street, Hilo, HI 96720, United States.
[5]Cancer Biology Program, University of Hawaii Cancer Center, 701 Ilalo Street, Honolulu, HI 96813, United States.
[6]These authors contributed equally to this work.

Correspondence and requests for materials should be addressed to Y. T. (yitang@ucla.edu), K. N. H. (houk@chem.ucla.edu), and J. Z. (jiahai@sioc.ac.cn)

One of the holy grails in chemical research is to design catalysts that effect specific reactions of complex molecules. Chemists rely on organo- or transition metal catalysts to control stereo-, regio-, and periselectivity (selectivity among possible pericyclic reactions). Nature achieves these types of selectivity with a variety of enzymes such as the recently discovered pericyclases – a family of enzymes that catalyze pericyclic reactions.[1] To date, the majority of characterized enzymatic pericyclic reactions are cycloadditions and it has been difficult to rationalize how observed selectivities are achieved.[2-12] We report here the discovery of two homologous groups of pericyclases that catalyze distinct reactions: one group catalyzes the first report of an Alder-ene reaction in biological systems; the second catalyzes a stereoselective hetero-Diels–Alder reaction. Guided by computational studies, we rationalized the observed differences in reactivities and designed mutants that reverse periselectivities from Alder-ene to hetero-Diels–Alder and *vice versa*. A combination of *in vitro* biochemical characterizations, computational studies, enzyme co-crystal structures, and mutational studies gave a complete picture of how high regio- and periselectivities are achieved in nearly identical active sites.

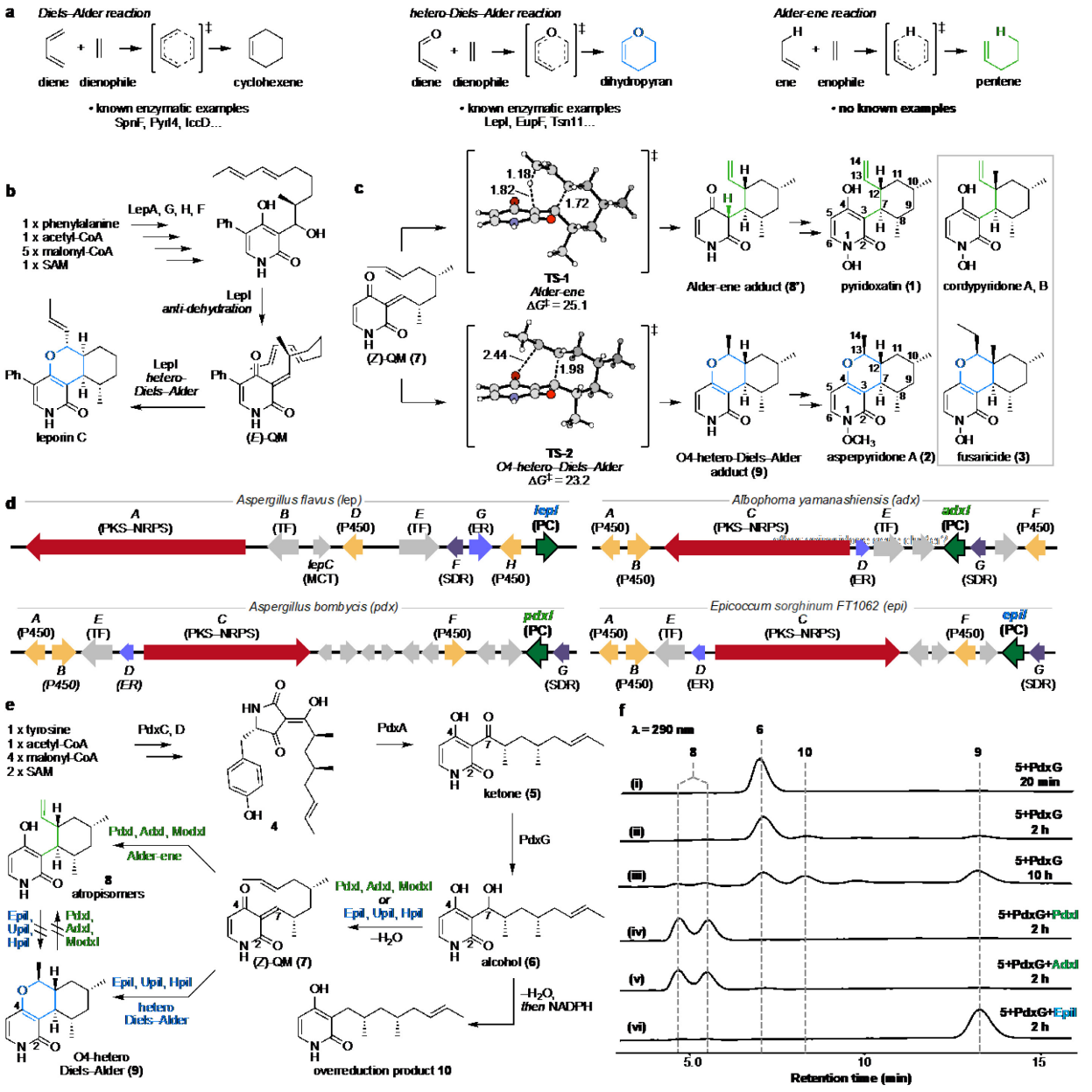


Figure 1 | Pericyclic reactions in natural product biosynthesis.
a, Known and unknown enzymatic examples of pericyclic reactions. **b**, The biosynthesis of leporin B highlighting a multifunctional *O*-methyltransferase-like pericyclase LepI. **c**, Theoretical investigations indicate that Alder-ene **TS-2** is nonenzymatically favored by 1.9 kcal·mol^{-1} over the hetero-Diels Alder **TS-1** from a common intermediate **7**. Further transformations lead to natural products pyridoxatin **1**, cordypyridones, asperpyridone **2** and fusaricide **3**. **d**, The biosynthetic gene cluster (*lep*) of leporin B from *Aspergillus flavus*, the putative biosynthetic gene cluster (*adx*) of pyridoxatin from *Albophoma yamanashiensis*, the putative biosynthetic gene cluster (*pdx*) of pyridoxatin from *Aspergillus bombycis*, and the putative biosynthetic gene cluster (*epi*) of fusaricide from *Epicoccum sorghinum* FT1062. PKS-NRPS, polyketide synthetase-non-ribosomal peptide synthetase; TF, transcription factor; MCT, monocarboxylate transporter; P450, cytochrome P450; SDR, short-chain dehydrogenase/reductase; ER, enoylreductase; PC, pericyclase. **e**, The proposed biosynthesis of the Alder-ene product (**8**) and the hetero-Diels–Alder product (**9**) from the common intermediate **7**. **f**, One-pot *in vitro* tandem assay of **5** with PdxG and in the presence or absence of selected pericyclases.

Pericyclic reactions are concerted chemical transformations in which all bonding changes occur in a cyclic array of atoms (**Fig. 1a**).[13] Diels–Alder and hetero-Diels–Alder reactions are classics in synthesis and pedagogy, and form cyclohexenes and pyrans, respectively.[14,15] Recently, these pericyclic reactions have been discovered to be catalyzed by enzymes in nature (**Fig. 1a**, **b**).[4] The Alder-ene reaction (**Fig. 1a**),[16] originally named the 'substituting addition' reaction by Kurt Alder's laboratory in 1943 and subject of his Nobel prize lecture,[17] is an efficient method for carbon-carbon bond formation and has been applied to total syntheses of complex polycyclic natural products.[18,19] While this reaction has been postulated to be involved in several biological processes,[20] there are no characterized examples of enzyme-catalyzed Alder-ene reactions in biology despite the frequency of substituted pentenes in natural products.[21] To this end, we set out to discover the enzymatic Alder-ene reaction in biosynthesis.

Chemical syntheses of racemic pyridoxatin (**1**)[22,23] and cordypyridone[24] and our previous work on related leporin 2-pyridone alkaloids[4,25] led us to hypothesize that the vinyl cyclohexane core of **1** could be formed by an Alder-ene reaction of a reactive (*E*)- or (*Z*)-quinone methide (QM) (**Fig. 1c**).[26] The activated QM can also undergo various hetero-Diels–Alder reactions involving either oxygen atom on the pyridone ring. To understand factors controlling the periselectivity, we expanded our study to include related natural products with the identical carbon backbones asperpyridone A (**2**)[27] and fusaricide (**3**)[28,29] that presumably derive from the hetero-Diels–Alder reaction of the same reactive QM. Because Alder-ene reactions are less exothermic (~–27 kcal·mol^{-1}) than cycloadditions (~–38 kcal·mol^{-1}), we anticipated that the Alder-ene reaction would be intrinsically more difficult than hetero-Diels–Alder reactions and periselectivity would strongly favor the latter.

To understand the reactivity of the QM in pericyclic reactions, we performed quantum mechanical calculations to determine transition state (TS) geometries and to quantify the barriers of possible pericyclic reactions from (*E*)- and (*Z*)-QM. Calculations indicate that the Alder-ene reaction has a slight preference (0.4 kcal·mol^{-1}) to occur from the (*Z*)- over the (*E*)-QM (**Extended Data Fig. 1**). From the (*Z*)-QM **7**, the Alder-ene reaction via **TS-1** is 1.9 kcal·mol^{-1} higher in energy than preferred hetero-Diels–Alder **TS-2** (**Fig. 1c** and **Extended Data Fig. 1**). This difference in Gibbs free energy corresponds to a nonenzymatic ratio of 1:9 Alder-ene to hetero-Diels–Alder adducts, and indicates that a dedicated pericyclase with great periselectivity must be involved in each of the biosynthetic pathways for the exclusive formation of either **1** or **2** (and **3**) from the producing strains.

To identify the potential pericyclases, we searched the genomes of **1**-producing strain *Albophoma yamanashiensis* and the related pyran-containing, **3**-producing strain *Epicoccum sorghinum* FT1062, as the **2**-producing strain was unavailable.[29,30] As shown in **Figure 1d**, the two genomes of *A. yamanashiensis* and *E. sorghinum* FT1062 both encode a homologous biosynthetic gene cluster (*adx* and *epi*, respectively) containing a polyketide synthase-nonribosomal peptide synthetase (*adxC*/*epiC*), a partnering enoylreductase (*adxD*/*epiD*), a ring expansion P450 (*adxA*/*epiA*), a putative *N*-hydroxylation P450 (*adxB*/*epiB*), a short-chain dehydrogenase/reductase (SDR) (*adxG*/*epiG*) and a putative *O*-methyltransferase (OMT) (*adxI*/*epiI*). We searched the National Center for Biotechnology Information (NCBI) database for other homologous clusters, which are conserved in many sequenced fungal strains such as *Aspergillus bombycis* (*pdx*), *Monosporascus cannonballus* (*modx*), *Uncinocarpus reesii* (*upi*) and *Hymenoscyphus scutula* (*hpi*) (**Fig. 1d** and **Extended Data Fig. 2**). Based on the sequence of previously discovered pericyclase LepI in leporin biosynthesis (**Fig. 1b, d**),[4] we hypothesized that

the predicted *O*-MT-fold enzymes in these pathways, which are AdxI, EpiI, PdxI, ModxI, UpiI and HpiI, are potential pericyclases. We proposed that each enzyme catalyzes the stereoselective dehydration of the alcohol (**6**) to **7**; AdxI from the **1**-producing strain would then catalyze the subsequent Alder-ene reaction; while EpiI from the **3**-producing strain would catalyze the hetero-Diels–Alder reaction (**Fig. 1e**). While the sequence identity between these enzymes are high (59–83%) (**Extended Data Fig. 2b**), they all display very low identity to LepI (~14%).

We performed the coupled *in vitro* reactions using enzymatically and chemoenzymatically prepared ketone **5** (supplementary information), the highly expressed and soluble SDR PdxG as the putative ketoreductase, and one of the proposed pericyclases. In the presence of PdxG and cofactor NADPH, **5** is reduced to the alcohol **6** (**Fig. 1f**, (i) and **Extended Data Fig. 3a**, **b**). In solution, **6** readily underwent nonenzymatic dehydration to generate both (*E*)- and (*Z*)-QM that rehydrate to form **6** and the C7 diastereomer (**Extended Data Fig. 4d**).[31] After 2 h, small amounts of O4-hetero-Diels–Alder product **9** and atropisomeric Alder-ene product **8** were detected in a ratio of 3 to 1 (**Fig. 1f**, (ii)) along with three other unidentified minor products (**Extended Data Figs. 1**, **4e**). Extended incubation times (10 h) with PdxG generated the overreduced **10**, which is presumably derived from a reduction of the QM (**Fig. 1f**, (iii) and **Extended Data Fig. 3c**). When AdxI, PdxI or ModxI was incubated with PdxG, NADPH and **5**, **8** was predominantly formed (>98:2, **8**:**9**) (**Fig. 1f**, (iv),(v) and **Extended Data Fig. 2c**). On the other hand, when we added EpiI or UpiI or HpiI to the reaction mixtures containing PdxG, NADPH, and **5**, the periselectivity was switched with **9** as the predominant product (<5:95, **8**:**9**) (**Fig. 1f**, (vi) and **Extended Data Fig. 2c**). Since **8** and **9** can be chemically interconverted,[24] we investigated whether these pericyclases could catalyze such a transformation. However, none of them were able to catalyze the interconversion (**Fig. 1e** and **Extended Data Fig. 4c**), which indicates these enzymes are dedicated

to catalyze the observed periselective reactions. Although these new pericyclases are predicted *S*-adenosyl-L-methionine (SAM)-dependent *O*-MTs,[4,32] none were copurifed with SAM (**Extended Data Fig. 4a**), nor was SAM or *S*-adenosyl-L-homocysteine (SAH) required for catalysis (**Extended Data Fig. 4b**). Overall, our biochemical data show that the group of AdxI, PdxI and ModxI represent the first enzymes characterized to catalyze the Alder-ene reaction forming a *bona fide* Alder-ene adduct. The group of EpiI, UpiI and HpiI were characterized as SAM-independent enzymes that catalyze a hetero-Diels–Alder reaction to form a *trans*-fused hexahydroisochromene in contrast to LepI that forms a *cis*-fused hexahydroisochromene. These discoveries expand the repertoire of reactions catalyzed in Nature.

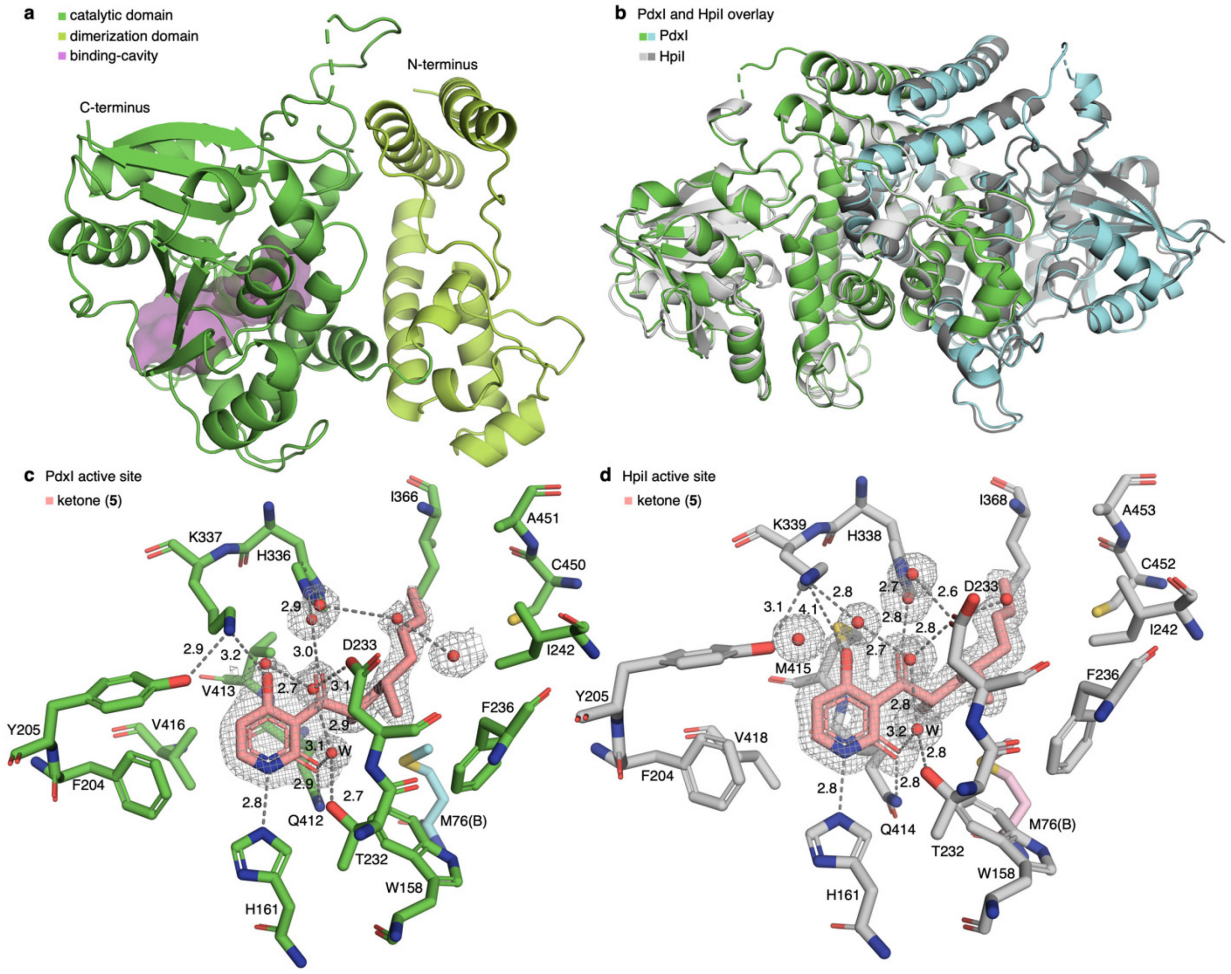


Figure 2 | Crystal structures of PdxI and HpiI. Simulated-annealing omit map shown in grey mesh and contoured at 1.0 σ. Hydrogen bond interactions are indicated with black dashed lines. **a**, Cartoon representation of *apo*-PdxI tertiary structure and binding-cavity (magenta). The *C*-terminal catalytic domain is shown in green and the *N*-terminal dimerization domain in lime. **b**, Overlap of interlocking homodimer structures of *apo*-PdxI and *apo*-HpiI. **c**, Active-site view of co-crystal structure of PdxI with substrate analogue ketone **5**. **d**, Active-site view of co-crystal structure of HpiI with substrate analogue ketone **5**. In **c** and **d**, M76 from Chain B is indicated in different colors.

To gain mechanistic insight into the enzyme-catalyzed Alder-ene reaction, the X-ray crystal structures of *apo*-PdxI, substrate analogue complex PdxI-**5**, and product complex PdxI-**8** were solved and refined to 2.0 Å, 2.0 Å, and 2.4 Å, respectively. The *apo*-PdxI structure adopts a classic α,β-Rossman *O*-MT fold (**Fig. 2a**)[25] and forms a woven dimer structure by interlocking the *N*-terminal helices (lime and royal blue, **Fig. 2b**). The overall *holo*-PdxI-**5** and PdxI-**8** structures

are very similar to *apo*-PdxI (root mean squared deviation (r.m.s.d.) of 0.162 and 0.202 for 831 Cα atoms, respectively). The chain A active site of PdxI-**5** and PdxI-**8** are shown in **Fig. 2c** and **Extended Data Fig. 5**. The pyridone ring forms hydrogen bonds with the side chains of K337, H161 and Q412, as well as water molecules via T232, D233 and H336 in the active site. The extended alkyl chain of **5** is pointed towards hydrophobic residue I366 and is not in a near attack conformation for the pericyclic reaction. The C7 ketone of substrate analogue **5** is positioned *syn* to the C4-hydroxyl and is indicative that the alcohol substrate undergoes a *syn*-dehydration facilitated by K337 to generate (*Z*)-QM **7** (**Extended Data Fig. 10**). In order to generate an (*E*)-QM the C7 alcohol would need to be rotated 180º and be *syn* to the C2-oxygen. The fact that this (*E*)-QM geometry differs greatly from that of the substrate analogue indicates that it is most probable that **7** is generated *in situ*. Substitution of K337 to alanine completely abolished the dehydration activity but can be rescued by mutation to arginine (**Extended Data Figs. 7**, **9a**), which supports the role of K337 as a general base for the dehydration step.

To investigate how the active site of PdxI catalyzes the energetically unfavorable Alder-ene reaction and suppresses the hetero-Diels–Alder reaction, we performed multiple 500 ns classical molecular dynamics (MD) simulations of **7** docked into PdxI. Since K337 is expected to be protonated after dehydration of **6** to **7**, we modeled side chain of K337 as an ammonium ion. We analyzed the conformation of the alkyl chain of the reactive **7** throughout the simulations and found that the alkyl chain can reorganize from the extended, unreactive conformation seen in PdxI-**5** into a reactive near attack conformation for 50-100 ns of the 500 ns simulation (**Extended Data Fig. 6c**). In all simulations, H161 and Q412 hydrogen bond to N1 and C2-carbonyl of the pyridone, respectively, whereas K337 and H336 hydrogen bond to the C4-carbonyl (**Extended Data Fig. 6a**, **b**). K337 maintains this hydrogen bond to the pyridone C4-carbonyl of **7** for longer durations

in the simulation, which implies the protonated K337 may facilitate the Alder-ene reaction by hydrogen bond catalysis.

Next, we quantified how the K337 and H161 hydrogen bonds affect the reaction rate with a truncated catalytic-residue 'theozyme' model; K337 is modeled as a methyl ammonium and H161 as an imidazole (**Fig. 3a**). The calculations indicate that the energetic barrier for the Alder-ene transition state in the theozyme (**TS-3**) is reduced by 11.7 kcal·mol^{-1}, a rate acceleration of >10^8, when compared to the nonenzymatic reaction (**TS-1**) (**Extended Data Fig. 1b**). Protonation of the carbonyl makes C7 highly electrophilic and decreases the nucleophilicity of the carbonyl oxygen. Both of these factors suppress the hetero-Diels–Alder transition state (**TS-4**) and favor **TS-3** by 3.4 kcal·mol^{-1}, a 5.8 kcal·mol^{-1} shift in $\Delta G^{\ddagger}$ relative to the predicted periselectivity of the nonenzymatic reactions (**Fig. 3a** and **Extended Data Fig. 1a**, **b**). Mechanistically, calculations indicate that **7** undergoes protonation by K337 and concomitant Alder-ene reaction via **TS-3** with an overall enthalpic barrier of 9.2 kcal·mol^{-1}. Whereas the O4-hetero-Diels–Alder via **TS-4** has a barrier of 12.6 kcal·mol^{-1}. This indicates that the PdxI active site alters the electronics of the reaction, by protonation of O4, to favor the Alder-ene reaction and achieve observed periselectivity.

It should be noted that the Alder-ene theozyme model also predicts a distinct, yet favorable hetero-Diels–Alder reaction via **TS-5** to form the *cis*-fused O2-hetero-Diels–Alder 4-pyridone adduct **11** (**Fig. 3a**). However, **11** is not found in the *in vitro* PdxI reaction mixtures (**Fig. 1f** and **Extended Data Figs. 4e**, **7**). This implies an additional degree of regio- and periselectivity exerted by PdxI to disfavor formation of **TS-5**. The PdxI structure with a docked **TS-5** indicates that the side chain of T232 and W causes disfavorable interactions with the terminal C14 methyl in **TS-5** (**Extended Data Fig. 5d**). The threonine residue acts as a steric wall to prevent **7** from accessing

the **TS-5** conformation. To test this hypothesis, we prepared the PdxI T232A/S/V mutants and performed the coupled *in vitro* reactions. Indeed, the less hindered T232A and T232S mutants generated an appreciable amount of a new compound as compared to the WT PdxI (**Fig. 3c** and **Extended Data Figs. 4e**, **7**). Large scale *in vitro* reaction of PdxI T232S mutant enabled the isolation and determination of the structure as the expected **11**, which is one of the nonenzymatic cyclized product from **7** (**Extended Data Fig. 1**, **4e**). Mutation to valine (PdxI T232V), the steric isostere of threonine, did not produce any **11** (**Fig. 3c** and **Extended Data Fig. 7**) supporting the role of T232 residue in sterically preventing the formation of **TS-5**.

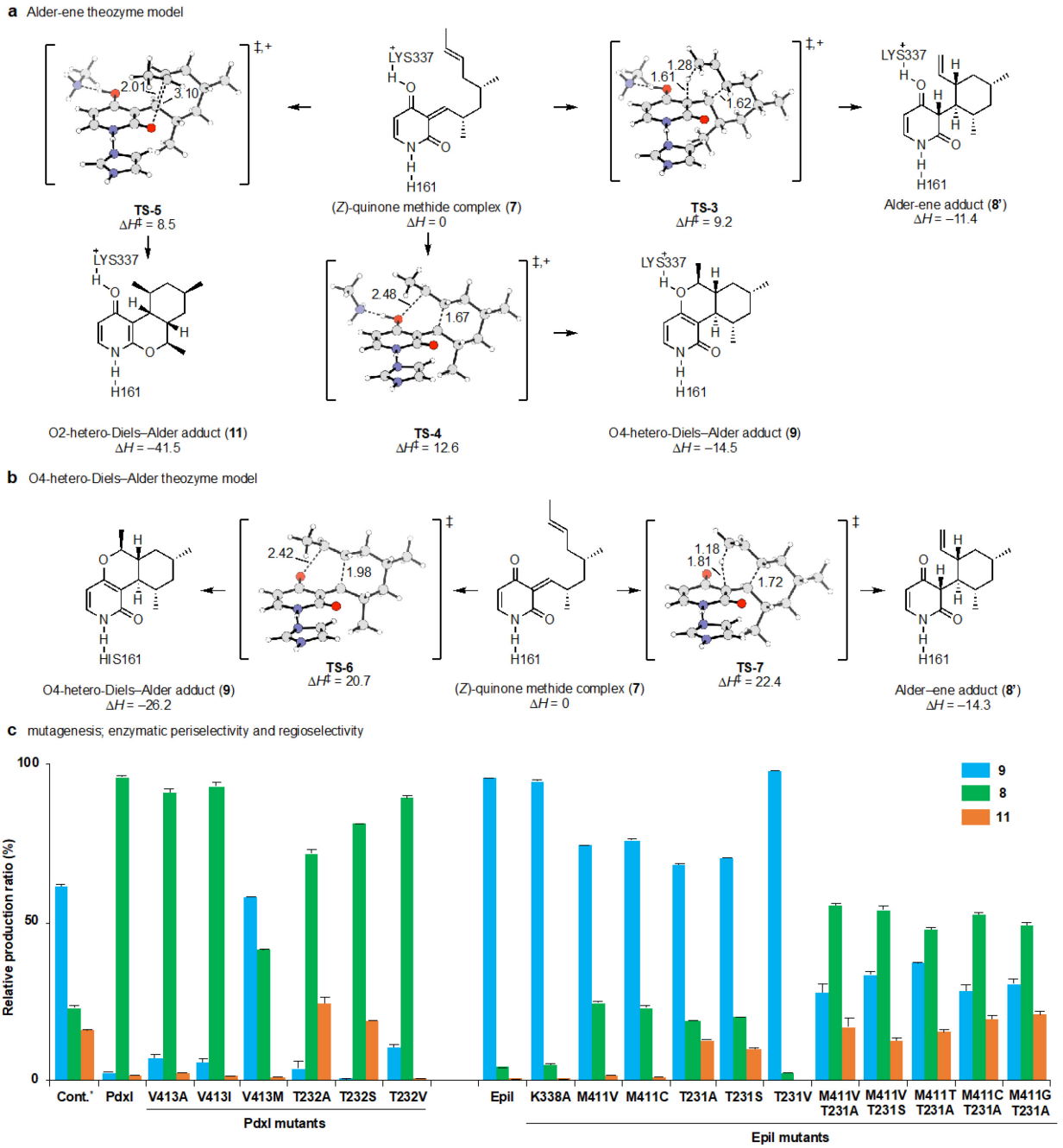


Figure 3 | Mechanism of periselective and regioselective pericyclic reactions.
a, Alder-ene theozyme model based on PdxI structure. **b**, O4-Hetero-Diels–Alder theozyme model confirmed by HpiI structure. **c**, Analysis of the relative production ratio of the O4-hetero-Diels–Alder adduct **9**, Alder-ene adduct **8**, and the O2-hetero-Diels–Alder adduct **11** from *in vitro* reaction of **5** with PdxG, NADPH, and selected pericyclases. To quantify the ratio, the reaction time of control sample without pericyclases (Cont.) was 12 hours, and with enzyme was 2 hours. Error bars indicate s.d. of three independent replicates. *Putative cyclized products other than **8**, **9**, and **11** were detected in the control reaction (**Extended Data Fig. 4e**).

PdxI crystal structures, molecular dynamics, and the 'theozyme' model indicate that the K337 residue acts as a general acid catalyst to favor the Alder-ene reaction over the O4-hetero-Diels–Alder reaction. Then, we calculated the O4-hetero-Diels–Alder 'theozyme' model (**Fig. 3b**) by removing the lysine from the PdxI K337 and H161 theozyme model (**Fig. 3a**), and indeed the O4-hetero-Diels–Alder reaction is preferred by 1.7 kcal·mol^{-1} (**TS-6** vs **TS-7**). This suggests that the group of enzymes (EpiI, HpiI and UpiI) that catalyze the O4-hetero-Diels–Alder reaction, in contrast to PdxI, must avoid proton transfer to the pyridone C4 carbonyl to achieve the opposite periselectivity. Nevertheless, the corresponding lysine is conserved in these three enzymes as well (**Extended Data Fig. 2b**). We solved and refined the crystal structures of *apo*-HpiI and HpiI-**5** to 1.3 Å and 1.5 Å resolution, respectively (**Fig. 2d**). These structures are highly similar to that of PdxI (r.m.s.d of 0.468 for 415 Cα atoms) (**Fig. 2b-d** and **Extended Data Fig. 5c**). Notably, the binding mode of **5** in PdxI and HpiI is essentially identical (**Figs. 2c**,**d** and **Extended Data Fig. 5c**) and nearly all amino acid residues in the active site between PdxI and HpiI are conserved within the two structures except for V413 in PdxI and the corresponding residue M415 in HpiI (**Fig. 2c**, **d** and **Extended Data Figs. 2b**, **5c**). This key residue sits below the pyridone binding site and neighbors the aforementioned lysine (K337 in PdxI, K338 in EpiI and K339 in HpiI). The HpiI-**5** complex clearly shows that K339, unlike K337 in PdxI, does not form a hydrogen bond to the 4-hydroxy on the pyridone ring; this distance is stretched out from 3.2 Å in PdxI to 4.1 Å in HpiI (**Fig. 2c**, **d**). To verify the lysine residue is not catalytic, we prepared the more stable mutant K338A of EpiI. In contrast to PdxI K337A mutant, EpiI K338A mutant retained the majority of the enzymatic activity (~80%) and showed the same periselectivity as WT EpiI (**Fig. 3c** and **Extended Data Figs. 8**, **9b**). Thus, the EpiI-catalyzed reactions do not require hydrogen bonding between the C4-oxygen on the pyridone ring and K338 (**Extended Data Figs. 8-10**). From the

crystal structure of HpiI, this loss of a hydrogen bond is caused by the bulkier side chain of M415, which shifts the lysine side chain further away from the substrate (**Fig. 2d** and **Extended Data Fig. 5c**). Consistently, homologous pericyclases such as EpiI, UpiI, and HpiI all catalyze the hetero-Diels–Alder reaction to form **9**, and the methionine residue is conserved (M411, M413, and M415, respectively). By contrast, pericyclases such as PdxI, AbxI and ModxI catalyze the Alder-ene reaction, and the valine residue (V413) is conserved (**Extended Data Fig. 2b**, **c**).

We next explored if the periselectivity of PdxI could be switched to favor the hetero-Diels–Alder reaction by the replacement of V413 residue with alanine, isoleucine or methionine. Mutating valine to alanine and isoleucine, residues smaller than methionine, retained the periselectivity of PdxI for the Alder-ene reaction (**Fig. 3c** and **Extended Data Figs. 7**, **9a**). The V413M mutant showed reversed periselectivity, switching the major reaction type from Alder-ene to hetero-Diels–Alder, forming **8** and **9** in a ratio of from >98:2 to 40:60 (**Fig. 3c** and **Extended Data Figs. 7**, **9a**). On the other hand, the mutation of M411 in EpiI to the less bulky valine or cysteine altered the product ratio (**8**:**9**) from 5:95 to 25:75 (**Fig. 3c** and **Extended data Figs. 8**, **9b**). The mutation of M411 alone in EpiI is not sufficient to reverse the periselectivity, suggesting that other factors such as shape complementarity to restrict the movement of **7** would also contribute to maintaining the observed periselectivity for the hetero-Diels–Alder reaction.

Since the conformational flexibility of **7** is expected to be restricted by both M411 and T231 residues in the EpiI active site based on the HpiI-**5** complex (**Fig. 2d**), we mutated the T231 residue to smaller alanine and serine residues. This mutation would increase the conformational flexibility of **7** in the EpiI active site to form the key hydrogen bonding with K338 for the Alder-ene reaction. Indeed, while the hetero-Diels–Alder product **9** still remains the major product, EpiI T231A and T231S mutants increased the ratio of Alder-ene product **8** along with the O2-hetero-

Diels–Alder product **11** as seen in PdxI T232A/S mutants (**Fig. 3c** and **Extended Data Figs. 8**, **9b**). Intriguingly, the double mutant M411V/T231A showed the reversed periselectivity with the ratio (**8**:**9**) of 66:33. Other less bulky double mutants such as M411V/T231S, M411T/T231A. M411C/T231A, and M411G/T231A showed similar periselectivies. Based on these results, we conclude that the replacement of the methionine and threonine with smaller residues enlarges the enzyme active site and allows for greater conformational sampling of **7** and protonation of the C4 carbonyl, thus leading to the opposite periselectivity.

Our results show that the group of PdxI, AdxI, and ModxI are multifunctional enzymes that catalyze the stereoselective *syn*-dehydration of **6** to **7** and the subsequent Alder-ene reaction of **7** to **8** in a stereo-, regio- and periselective manner. In contrast, the group of EpiI, UpiI, and HpiI catalyze the same stereoselective *syn*-dehydration of **6** to **7** but with orthogonal periselectivity and catalyze the hetero-Diels–Alder reaction of **7** to **9**. Computational studies, comparative analysis of the enzyme-cocrystal structures, and site-directed mutagenesis provided a detailed picture of the catalytic mechanism for PdxI and EpiI (**Extended Data Fig. 10**). PdxI utilizes K337 as general acid catalyst to facilitate the otherwise energetically unfavorable Alder-ene reaction, while the methionine substitution in EpiI abolishes this interaction to allow only the O4-hetero-Diels–Alder reaction. The steric effect of T232 in PdxI and T231 in EpiI inhibits the formation of the O2-hetero-Diels–Alder product **11** to further control regioselectivity.

In conclusion, we have characterized two homologous groups of enzymes and identified how subtle evolutionary divergence leads to the production of different natural products. The insight gained from our research serves as a basis for developing new biocatalysts that catalyze various natural and unnatural Alder-ene and hetero-Diels–Alder reactions that are valuable synthetic transformations.

References

1. Jamieson, C. S., 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.* **36**, 698–713 (2019).
2. Kim, H. J., Ruszczycky, M. W., Choi, S., Liu, Y. & Liu, H. Enzyme-catalysed [4+2] cycloaddition is a key step in the biosynthesis of spinosyn A. *Nature* **473**, 109 (2011).
3. Wever, W. J. *et al.* Chemoenzymatic Synthesis of Thiazolyl Peptide Natural Products Featuring an Enzyme-Catalyzed Formal [4 + 2] Cycloaddition. *J. Am. Chem. Soc.* **137**, 3494–3497 (2015).
4. Ohashi, M. *et al.* SAM-dependent enzyme-catalysed pericyclic reactions in natural product biosynthesis. *Nature* **549**, 502 (2017).
5. Bailey, S. S. *et al.* Enzymatic control of cycloadduct conformation ensures reversible 1,3-dipolar cycloaddition in a prFMN-dependent decarboxylase. *Nat. Chem.* **11**, 1049–1057 (2019).
6. Chen, Q. *et al.* Enzymatic Intermolecular Hetero-Diels-Alder Reaction in the Biosynthesis of Tropolonic Sesquiterpenes. *J. Am. Chem. Soc.* **141**, 14052–14056 (2019).
7. Zhang, B. *et al.* Enzyme-catalysed [6+4] cycloadditions in the biosynthesis of natural products. *Nature* **568**, 122–126 (2019).
8. Little, R. *et al.* Unexpected enzyme-catalysed [4+2] cycloaddition and rearrangement in polyether antibiotic biosynthesis. *Nat. Catal.* (2019). doi:10.1038/s41929-019-0351-2
9. Dan, Q. *et al.* Fungal indole alkaloid biogenesis through evolution of a bifunctional reductase/Diels–Alderase. *Nat. Chem.* **11**, 972–980 (2019).
10. Zhang, Z. *et al.* Enzyme-Catalyzed Inverse-Electron Demand Diels–Alder Reaction in the Biosynthesis of Antifungal Ilicicolin H. *J. Am. Chem. Soc.* **141**, 5659–5663 (2019).
11. Gustin, D. J. *et al.* Heavy Atom Isotope Effects Reveal a Highly Polarized Transition State for Chorismate Mutase. *J. Am. Chem. Soc.* **121**, 1756–1757 (1999).
12. DeClue, M. S., Baldridge, K. K., Künzler, D. E., Kast, P. & Hilvert, D. Isochorismate Pyruvate Lyase: A Pericyclic Reaction Mechanism? *J. Am. Chem. Soc.* **127**, 15002–15003 (2005).
13. Woodward, R. B. & Hoffmann, R. The Conservation of Orbital Symmetry. *Angew. Chemie Int. Ed. English* **8**, 781–853 (1969).
14. Nicolaou, K. C., Snyder, S. A., Montagnon, T. & Vassilikogiannakis, G. The Diels–Alder Reaction in Total Synthesis. *Angew. Chemie Int. Ed.* **41**, 1668–1698 (2002).
15. Corey, E. J. & Cheng, X. M. *The Logic of Chemical Synthesis*. (Wiley, 1989).
16. Hoffmann, H. M. R. The Ene Reaction. *Angew. Chemie Int. Ed. English* **8**, 556–577 (1969).
17. Alder, K. Nobel Lectures, Chemistry 1942–1962. in 267–303 (Elsevier Publishing Company, 1964).
18. Niu, D. & Hoye, T. R. The aromatic ene reaction. *Nat. Chem.* **6**, 34–40 (2014).
19. Mikami, K. & Shimizu, M. Asymmetric ene reactions in organic synthesis. *Chem. Rev.* **92**, 1021–1050 (1992).
20. Jensen, A. W., Mohanty, D. K. & Dilling, W. L. The growing relevance of biological ene

reactions. *Bioorg. Med. Chem.* **27**, 686–691 (2019).

21. Lin, C.-I., McCarty, R. M. & Liu, H. The Enzymology of Organic Transformations: A Survey of Name Reactions in Biological Systems. *Angew. Chemie Int. Ed.* **56**, 3446–3489 (2017).
22. Snider, B. B. & Lu, Q. Total Synthesis of (.+-.)-Pyridoxatin. *J. Org. Chem.* **59**, 8065–8070 (1994).
23. Snider, B. B. & Qing, L. A two-step synthesis of pyridoxatin analogues. *Tetrahedron Lett.* **35**, 531–534 (1994).
24. Jones, I. L., Moore, F. K. & Chai, C. L. L. Total Synthesis of (±)-Cordypyridones A and B and Related Epimers. *Org. Lett.* **11**, 5526–5529 (2009).
25. Cai, Y. *et al.* Structural basis for stereoselective dehydration and hydrogen-bonding catalysis by the SAM-dependent pericyclase LepI. *Nat. Chem.* **11**, 812–820 (2019).
26. Appendino, G., Cravotto, G., Toma, L., Annunziata, R. & Palmisano, G. The Chemistry of Coumarin Derivatives. Part VI. Diels-Alder Trapping of 3-Methylene-2,4-chromandione. A New Entry to Substituted Pyrano[3,2-c]coumarins. *J. Org. Chem.* **59**, 5556–5564 (1994).
27. Qiao, Y. *et al.* Asperpyridone A: An Unusual Pyridone Alkaloid Exerts Hypoglycemic Activity through the Insulin Signaling Pathway. *J. Nat. Prod.* **82**, 2925–2930 (2019).
28. McBrien, K. D. *et al.* Fusaricide, a New Cytotoxic N-Hydroxypyridone from Fusarium sp. *J. Nat. Prod.* **59**, 1151–1153 (1996).
29. Li, C., Sarotti, A. M., Yang, B., Turkson, J. & Cao, S. A New N-methoxypyridone from the Co-Cultivation of Hawaiian Endophytic Fungi Camporesia sambuci FT1061 and Epicoccum sorghinum FT1062. *Molecules (Basel, Switzerland)* **22**, (2017).
30. Lee, H. J. *et al.* Pyridoxatin, an inhibitor of gelatinase A with cytotoxic activity. *J. Microbiol. Biotechnol.* **6**, 445–450 (1996).
31. Singh, M. S., Nagaraju, A., Anand, N. & Chowdhury, S. ortho-Quinone methide (o-QM): a highly reactive, ephemeral and versatile intermediate in organic synthesis. *RSC Adv.* **4**, 55924–55959 (2014).
32. Singh, S. *et al.* Structural characterization of the mitomycin 7-O-methyltransferase. *Proteins Struct. Funct. Bioinforma.* **79**, 2181–2188 (2011).

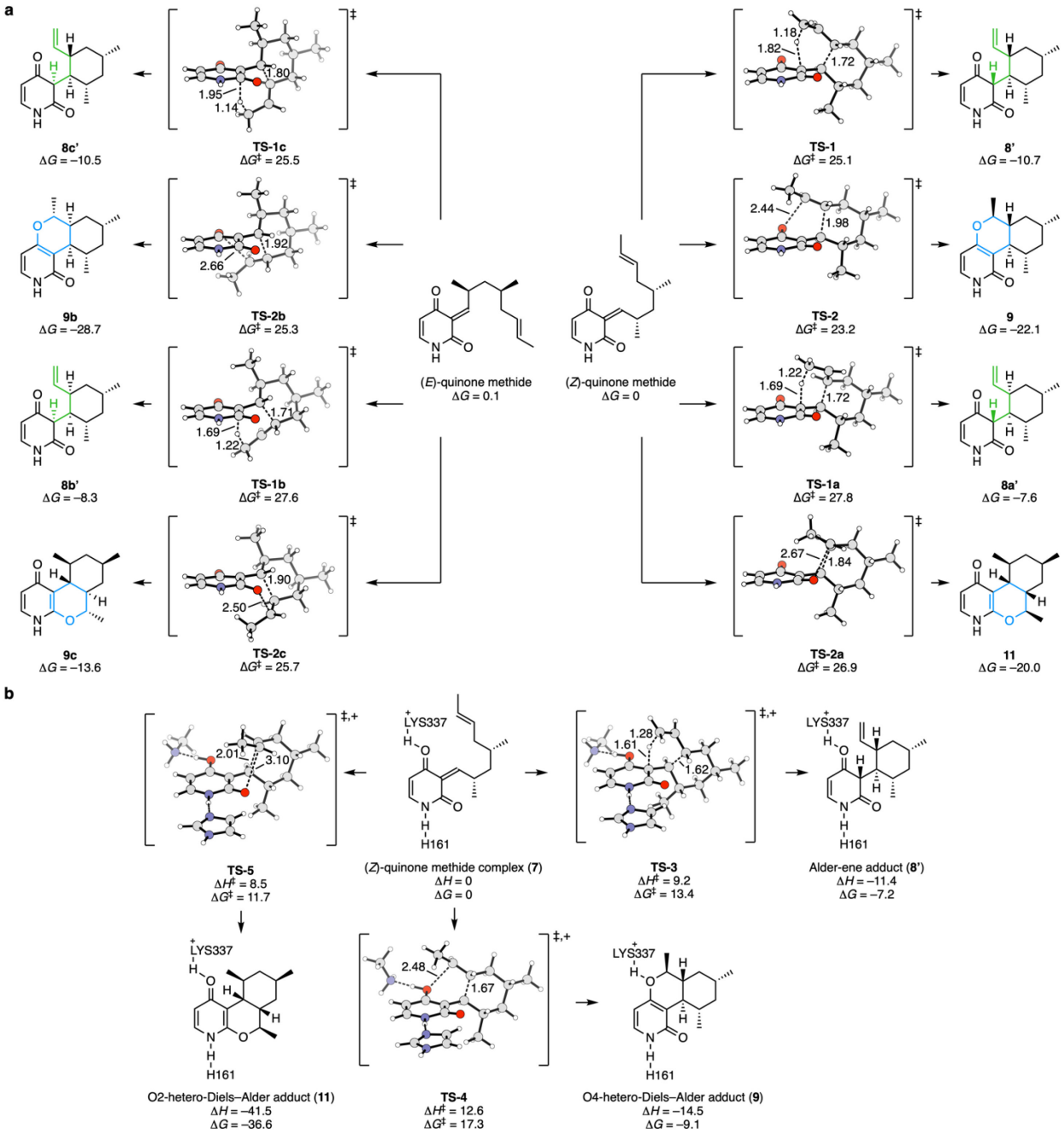
Supplementary Information is available in the online version of the paper.

Acknowledgement This work was supported by the NIH (1R01AI141481) to Y.T, K.N.H, and N.G., the NSF (CHE-1361104) to K.N.H., NSFC (91856202) to J.Z., and SMSTC (18430721500, 19XD1404800) to J.Z. The authors thank the staff of beamlines BL17U1, BL18U1 and BL19U1 of Shanghai Synchrotron Radiation Facility for access and help with the X-ray data collection. Y.C. thanks J. Gan for assistance with structure refinement.

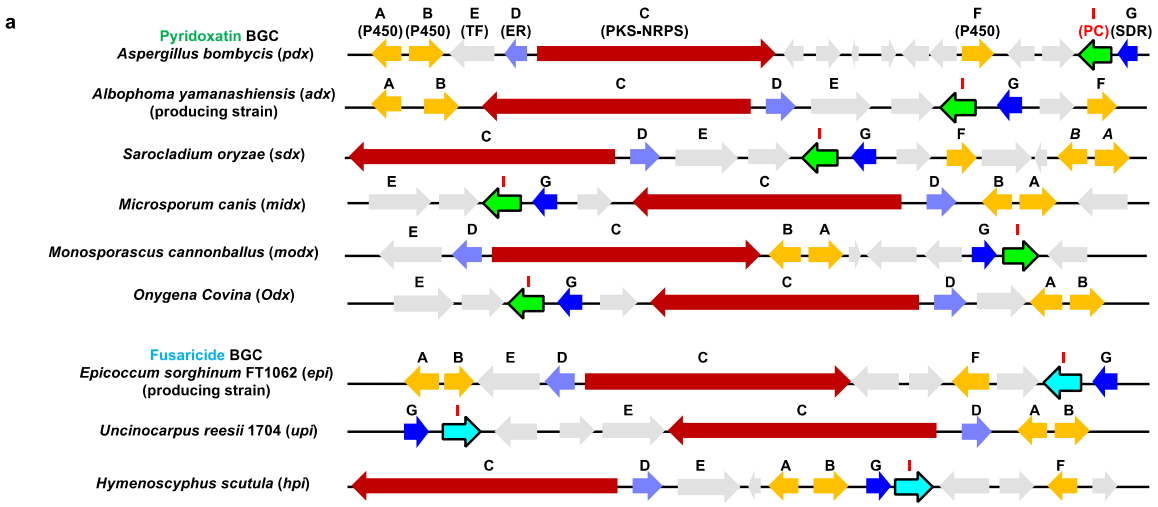
Author contributions M.O., C.S.J., K.N.H. and Y.T. developed the hypothesis and conceived the idea for the study. M.O., C.S.J., Y.C., J.Z., K.N.H. and Y.T. designed the experiments. M.O. performed all *in vivo* and *in vitro* experiments, as well as compound isolation and characterization. D.T., D.K., and M.-C.T. performed compound isolation and characterization. S.M.A., J.V.C., J.S.B., and E.P. performed synthesis of compounds. M.O. and T.B.K. performed bioinformatic analysis to identify the biosynthetic gene cluster. M.O. and Y.C. performed protein purification. Y.C. preformed all structural biology. C.S.J. performed all computational experiments. All authors analyzed and discussed the results. M.O., C.S.J., Y.C., J.Z., K.N.H. and Y.T. prepared the manuscript. M.O., C.S.J., and Y.C. contributed equally to this work.

Competing financial interests The authors declare no competing financial interests.

Extended Data Figures (10/10)

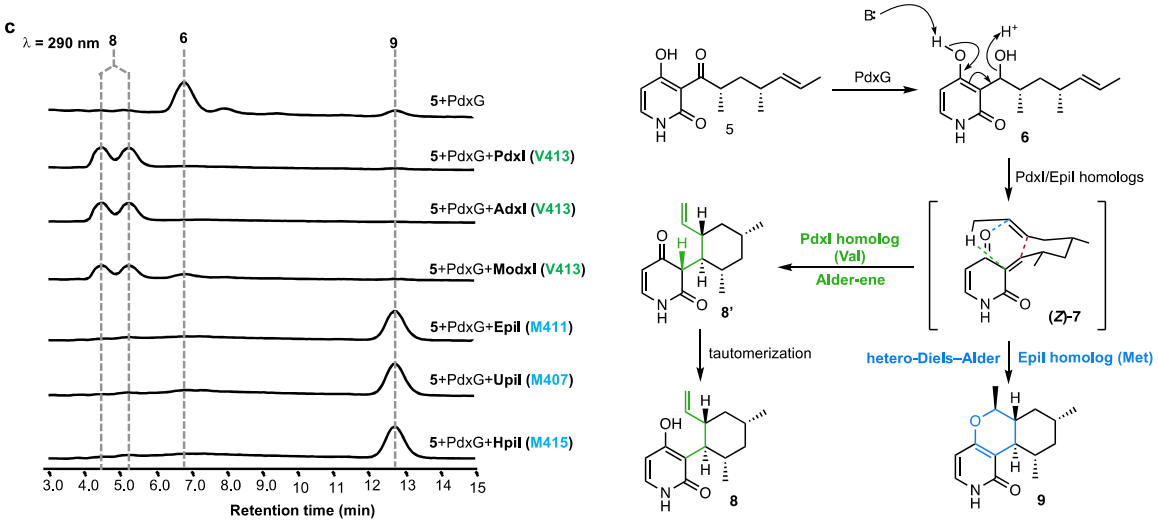


Extended Data Figure 1. Density functional theory calculations for non-enzymatic Alder-ene and hetero-Diels–Alder reactions from (*Z*)-QM and (*E*)-QM and PdxI theozyme. (**a**) Transition states, products, and energies for eight hetero-Diels–Alder and Alder-ene reactions are shown. In the transition states, the Alder-ene reactions adopt a conformation where the pyridone and forming cyclohexane are perpendicular to each other compared to the hetero-Diels–Alder reactions that are more co-planar in geometry. The Alder-ene reactions are synchronous and the hetero-Diels–Alder reactions are asynchronous, but concerted. **TS-1** and **TS-2** lead to **8'** and **9** with barriers of 23.2 and 25.1 kcal·mol^{-1}. The structures **8a'**, **b'**, **c'** and **9a**, **b**, **c** are isomers of natural product scaffolds with barriers greater than **TS-1**. (**b**) Alder-ene theozyme of (*Z*)-quinone methide complex leading to Alder-ene adduct (**8'**), O4- and O2-hetero-Diels–Alder adducts (**9** and **11**) with energies reported as enthalpies and Gibbs free energies.



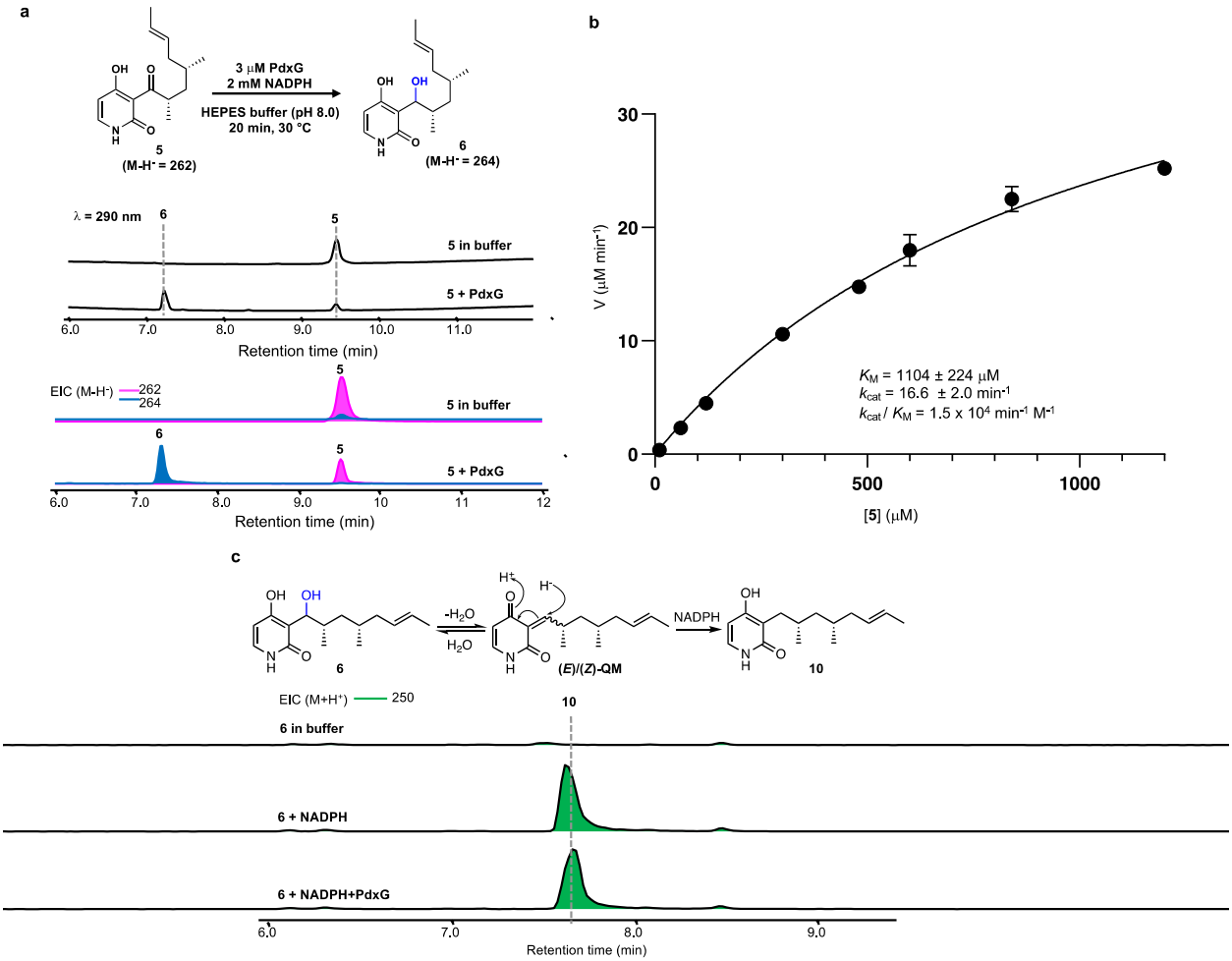
b

	PdxI Identity	EpiI Identity	Amino acid residues forming the active site																
			158	161	204	205	229	232	233	236	242	333	336	337	366	409	412	413	416
PdxI	100%	66%	W	H	F	Y	L	T	D	F	I	T	H	K	I	I	Q	V	V
SdxI	73%	61%	W158	H161	F204	Y205	L229	T232	D233	F236	I242	T333	H336	K337	I366	I409	Q412	V413	V416
AdxI	83%	64%	W158	H161	F204	Y205	L229	T232	D233	F236	I242	T333	H336	K337	I366	I409	Q412	V413	V416
ModxI	78%	61%	W158	H161	F204	Y205	L229	T232	D233	F236	V242	T333	H336	K337	I366	I409	Q412	V413	V416
MidxI	72%	59%	W158	H161	F204	Y205	L229	T232	D233	F236	I242	T333	H336	K337	I366	I408	Q411	V412	V415
OdxI	70%	60%	W158	H161	F204	Y205	L229	T232	D233	F236	V242	T333	H336	K337	I366	L409	Q412	L413	V416
EpiI	66%	100%	W158	H161	F204	Y205	L228	T231	D232	F235	V243	T334	H337	K338	I367	I407	Q410	M411	L414
UpiI	77%	63%	W158	H161	F204	Y205	L229	T232	D233	F236	M242	T333	H336	K337	I366	I403	Q406	M407	V410
HpiI	79%	67%	W158	H161	F204	Y205	L229	T232	D233	F236	I242	T335	H338	K339	I368	A411	Q414	M415	V419



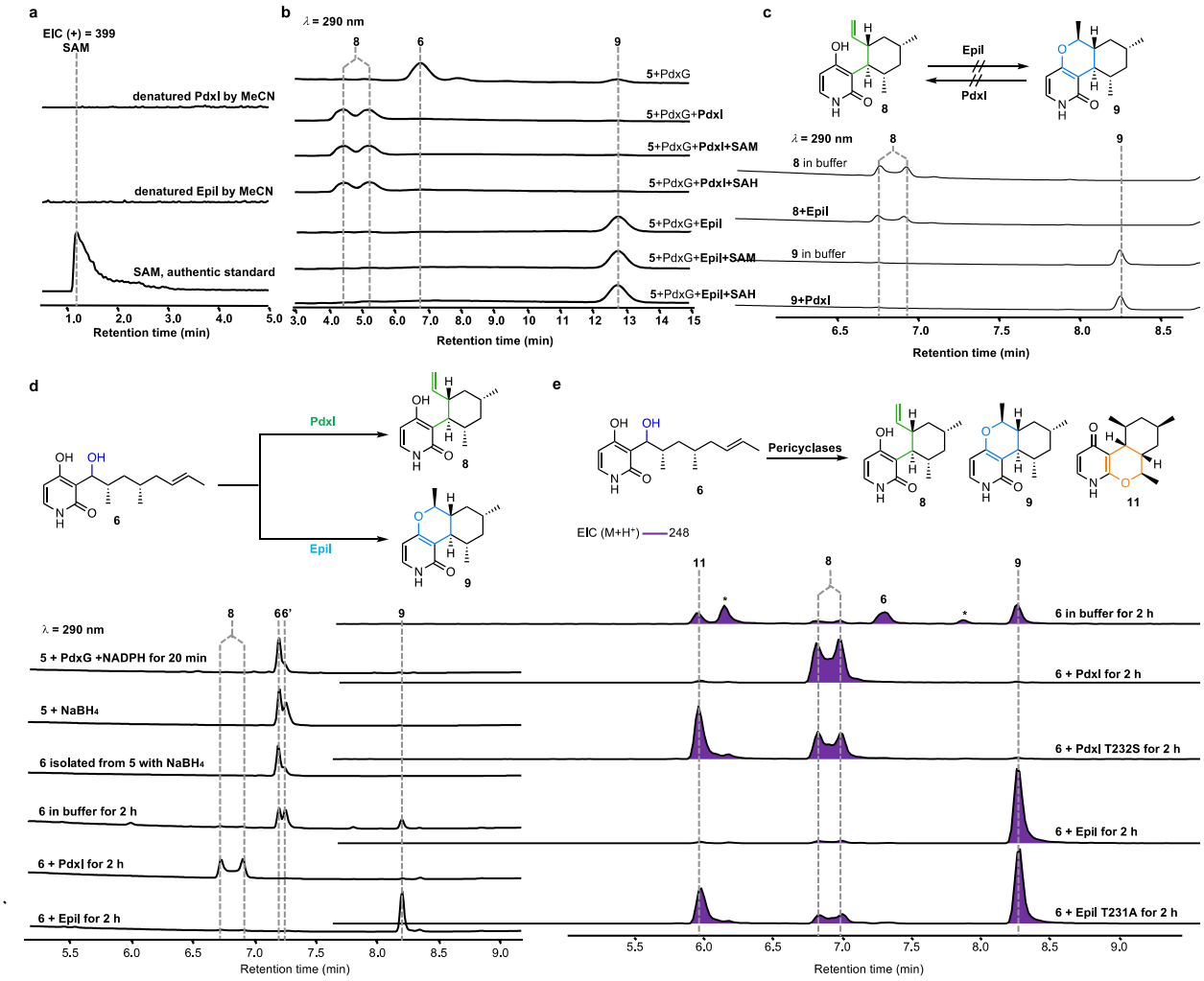
Extended Data Figure 2. Homologous biosynthetic gene clusters of pyridoxatin (1) and fusaricide (3) and the function of PdxI and EpiI homologs.

(**a**) Putative biosynthetic gene clusters of **1** and **3** (and **2**), and their homologous biosynthetic gene cluster found in NCBI database. (**b**) Key active site residues shown in an alignment with those from PdxI and EpiI homologs. Key residues for the PdxI and EpiI catalysis are colored. (**c**) *In vitro* analysis of PdxG and selected pericyclases using **5** as the starting substrate. The periselectivity can be correlated with the identity of the amino acid at position 413 (in PdxI, indicated in red dashed box). If valine occupied this position, the enzyme is shown to catalyze the Alder-ene reaction. On the other hand, if methionine occupies the position, the enzyme is shown to catalyze the hetero-Diels–Alder reaction.



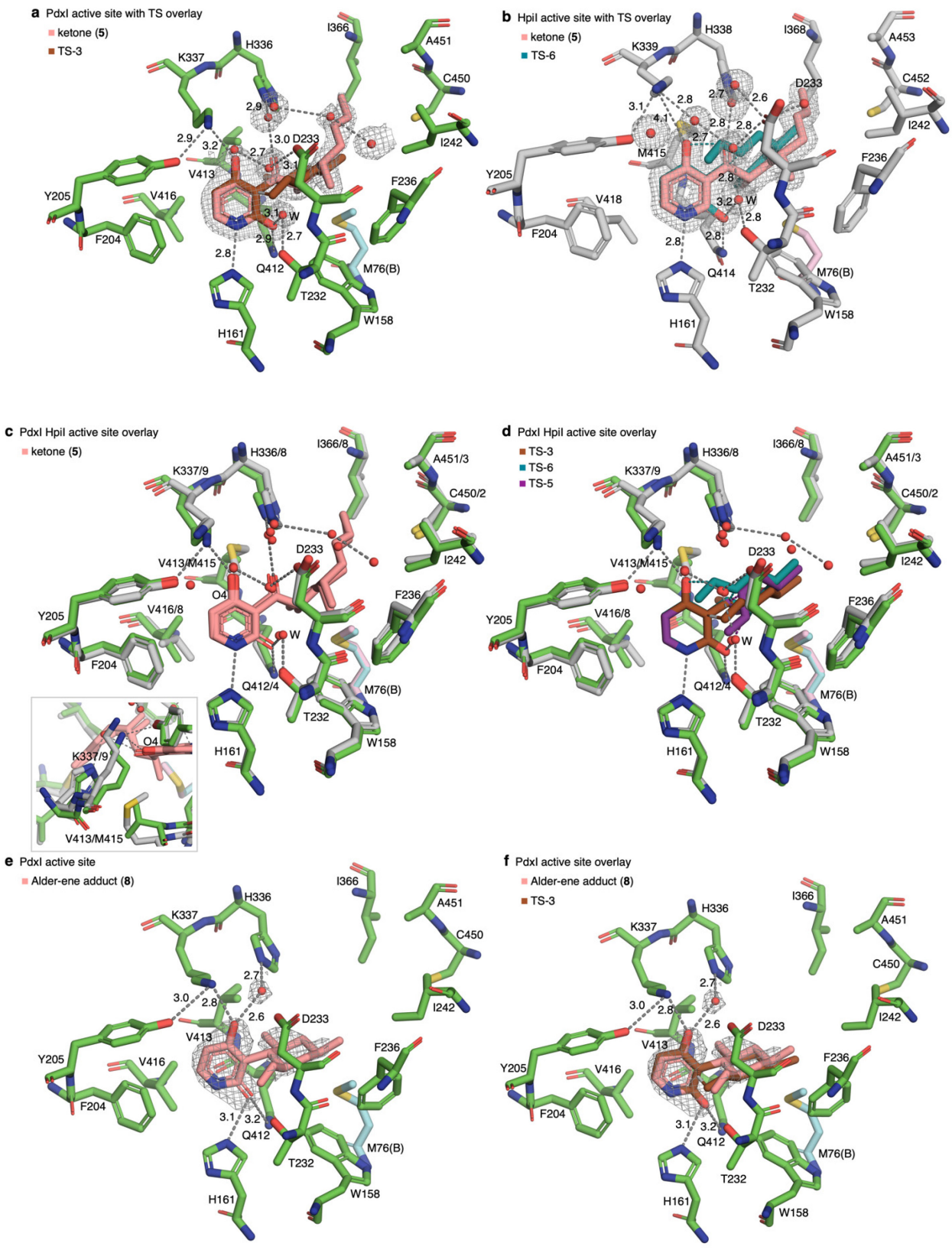
Extended Data Figure 3. Biochemical characterization of the ketoreductase PdxG.

(**a**) *In vitro* reaction of 3 μM PdxG with 2 mM NADPH using 600 μM **5** as the substrate. 60% conversion from **5** to **6** was observed within 20 min. (**b**) Kinetic analysis of PdxG-catalyzed reduction of **5**. Reaction mixtures containing 3 μM PdxG, 2 mM NADPH and different concentration of **5** (10 μM to 1.2 mM) were incubated at 30 °C for 20 min. Error bars indicate s.d. of three independent replicates. (**c**) Formation of **10** from **6** in the presence of 2 mM NADPH can be observed both with and without PdxG. Compound **6** was obtained from chemical reduction of **5** with $NaBH_4$. Since **10** can be formed in the presence of NADPH only, we conclude NADPH can nonenzymatically reduce the QM to **10**, which explain the result in Fig. 1f.



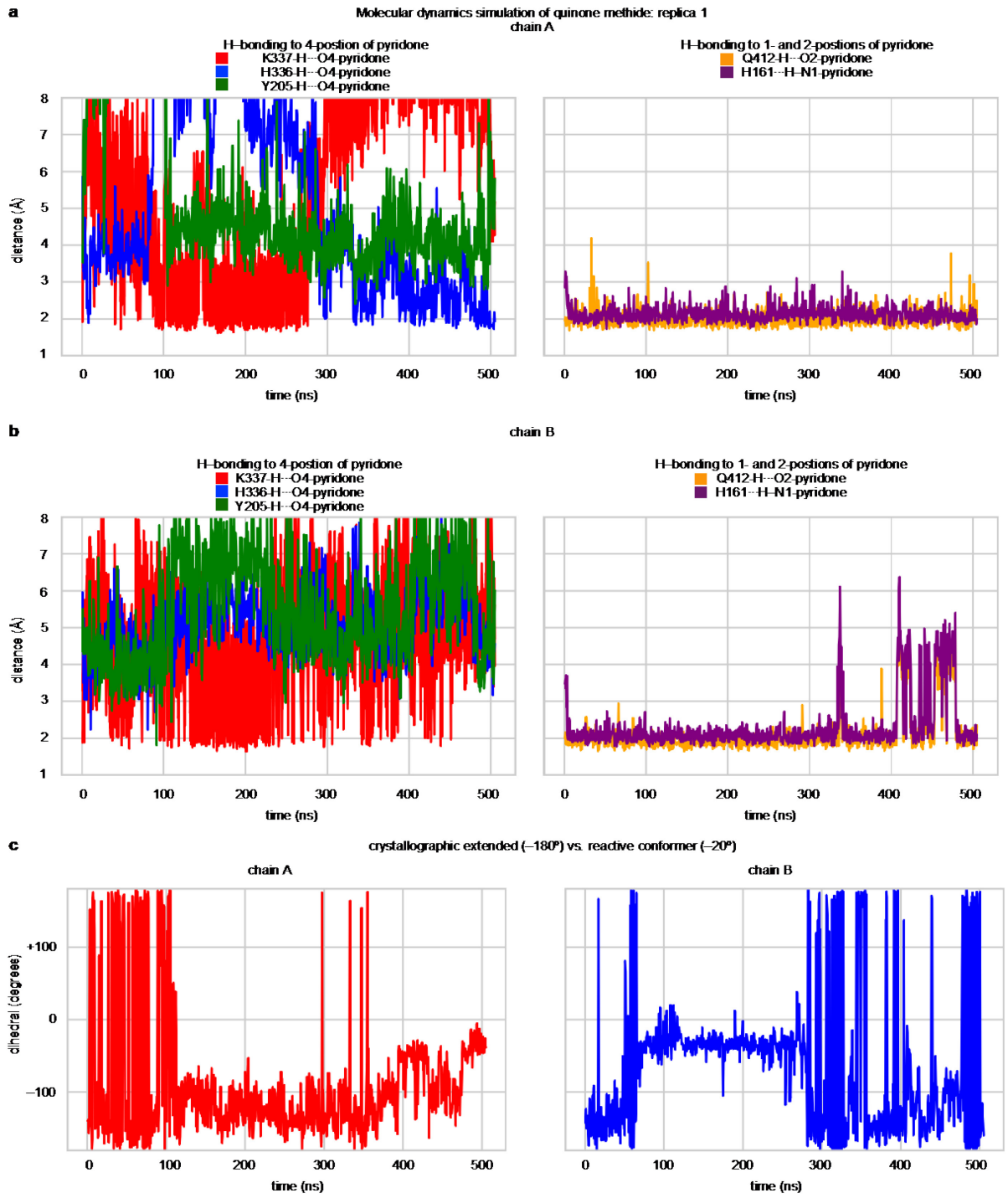
Extended Data Figure 4. Biochemical characterization of PdxI and EpiI.

(**a**) LC/MS analysis of denatured PdxI and EpiI shows no trace SAM after purification. (**b**) HPLC analysis of *in vitro* reaction of 150 μM **5** with 3 μM PdxG, 1 mM NADPH and 30 μM PdxI or 20 μM EpiI at 30 °C for 2 h in the presence and absence of cofactors. SAM or SAH do not affect the enzymatic activity of PdxI and EpiI. (**c**) Since the interconversion of **8** and **9** could be envisioned by hydroalkoxylation/retro-hydroalkoxylation, we examined the possibility that PdxI and EpiI catalyze the reaction of **9** to **8** and **8** to **9**, respectively. Thus, we performed the analysis of the *in vitro* reaction of 100 μM **9** or **8** with 30 μM PdxI or 20 μM EpiI at 30 °C for 24 h. However, no conversion of **9** to **8** and **8** to **9** was observed. (**d**) *In vitro* reaction of PdxI and EpiI using **6** as the substrate. To obtain **6** for in vitro reaction, we chemically reduced **5** by $NaBH_4$. Since this reduction proceeds non-stereoselectively, **6** and diastereomer **6′** were formed. After the isolation of **6** and **6′** by HPLC, the fractions containing **6** were not concentrated because of the instability and were immediately used as the substrate for PdxI and EpiI. (**e**) LC/MS analysis of *in vitro* reactions catalyzed by pericyclases using **6** as the substrate. Shown are compounds detected by selected ion monitoring at $(M+H)^+$ of 248. In this mode, **6** is detected as the fragment ion. In the absence of either enzyme, **6** was converted to several products nonenzymatically, including **11**. Minor compounds not isolated are indicated with *. In the presence of PdxI, **6** was nearly all converted to **8**. In the presence of EpiI, **6** was nearly all converted to **9**. Mutation of T232S in PdxI or T231A in EpiI changed selectivities of the enzymes to give other products, including **11**. For additional mutagenesis data, please see **Extended Data Figs. 7-8.**



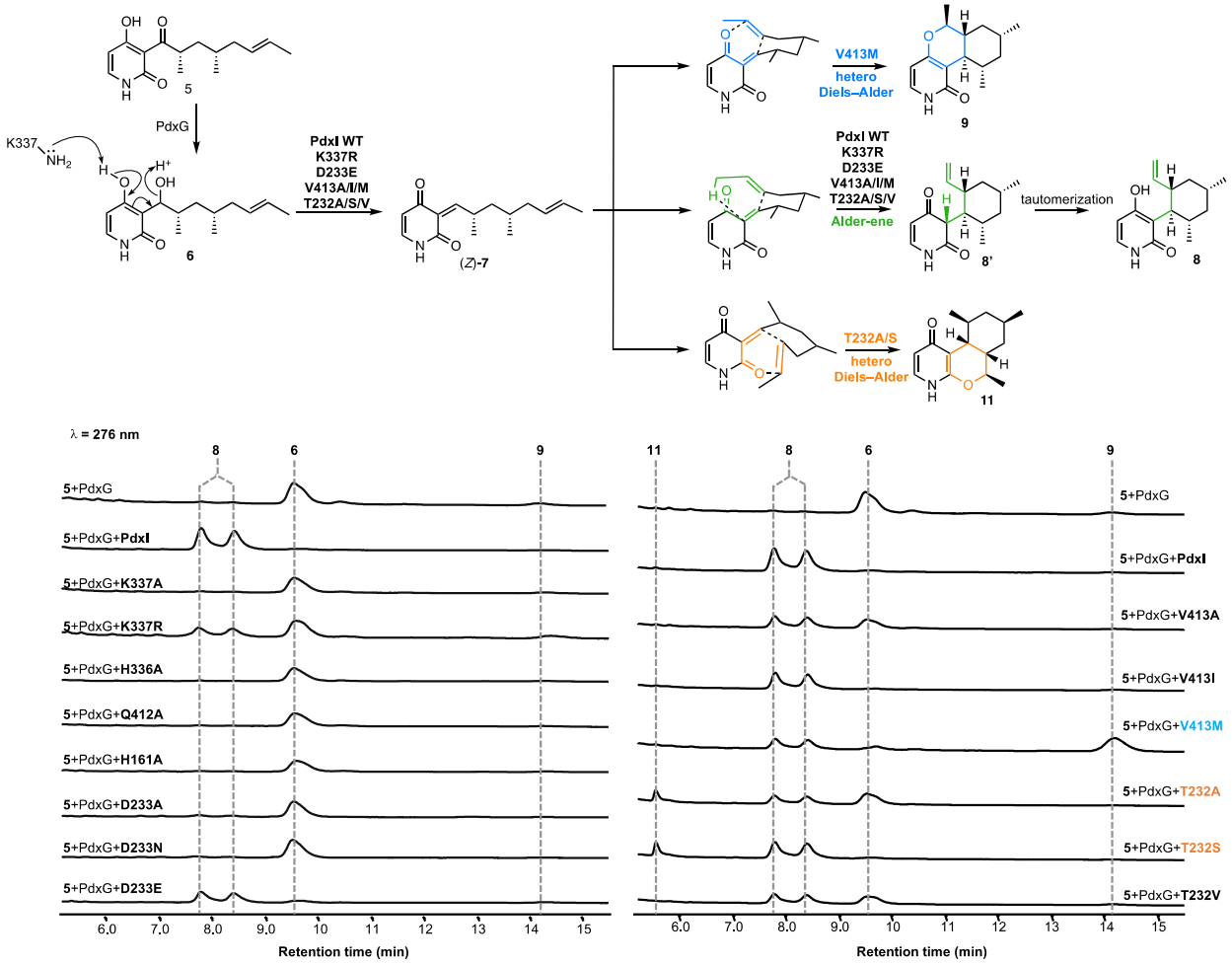
Extended Data Figure 5. Overlays of crystal structures with transition state structures for Alder-ene and hetero-Diels–Alder reactions. (**a**) Overlay of Alder-ene **TS-3** with **5** bound in PdxI. Note

the extended conformation of the alkyl chain versus the folded transition state geometry. The pyridone is bound by hydrogen bonds from K337, H161, Q412, and water mediated hydrogen bonds from T232, D233 and H336. (**b**) Overlay O4-hetero-Diels–Alder **TS-6** with **5** bound in HpiI. The pyridone is bound by hydrogen bonds from H161, Q414, and water mediated hydrogen bonds from T232, D233 and H338. Note that the K339 hydrogen bond to the pyridone O4 is not present in this structure. (**c**) Overlay of PdxI-**5** and HpiI-**5**. Omit maps not shown for clarity. Both PdxI and HpiI bind the pyridone such that it is prone to a *syn*-dehydration assisted by K337 (PdxI) or water molecule W (HpiI) and water molecules surrounding C7. The inset shows how V413 (in PdxI) or M415 (in HpiI) affects the orientation of the lysine residue (K337 in PdxI or K339 in HpiI) and its ability to hydrogen bond to the 4-OH of the pyridone. (**d**) Overlay of PdxI-**5**, HpiI-**5**, Alder-ene **TS-3**, and O4-hetero-Diels–Alder **TS-6**, and O2-hetero-Diels–Alder **TS-5**. Omit maps not shown for clarity. **TS-3** and **TS-6** bind in the active site sans disfavorable interactions whereas **TS-5** clashes with T232. As both the calculated Alder-ene transition structure **TS-3** and hetero-Diels–Alder **TS-6** are quite similar in geometry and both easily fit into the PdxI active site, PdxI cannot solely rely on shape complementarity to catalyze the reaction with observed periselectivity. (**e**) Chain B active site of PdxI-product (**8**) complex. Note the closer distances between K337 and the pyridone O4, the change in coordination of water mediated hydrogen bond from H336, and H161 shifting from a N1 hydrogen bond (in PdxI-**5**) to an O2 hydrogen bond. (**f**) Overlay of Alder-ene **TS-3** with **8** bound in PdxI. Note high similarity in structures of **8** and **TS-3**. This suggests that the enzyme distorts the product structure towards that of the Alder-ene transition state.

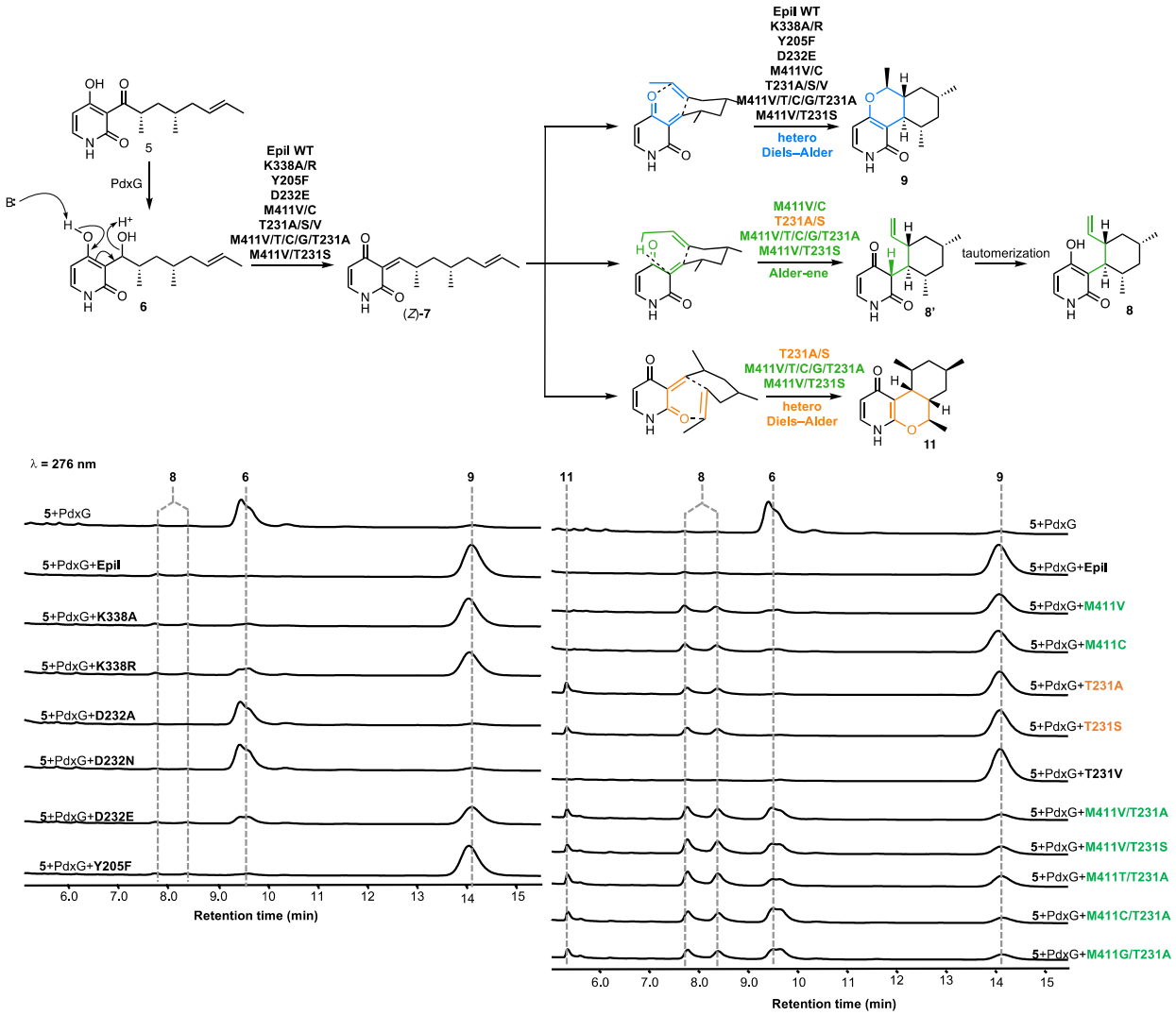


Extended Data Figure 6. Molecular dynamic simulation of 7 in the PdxI active site.

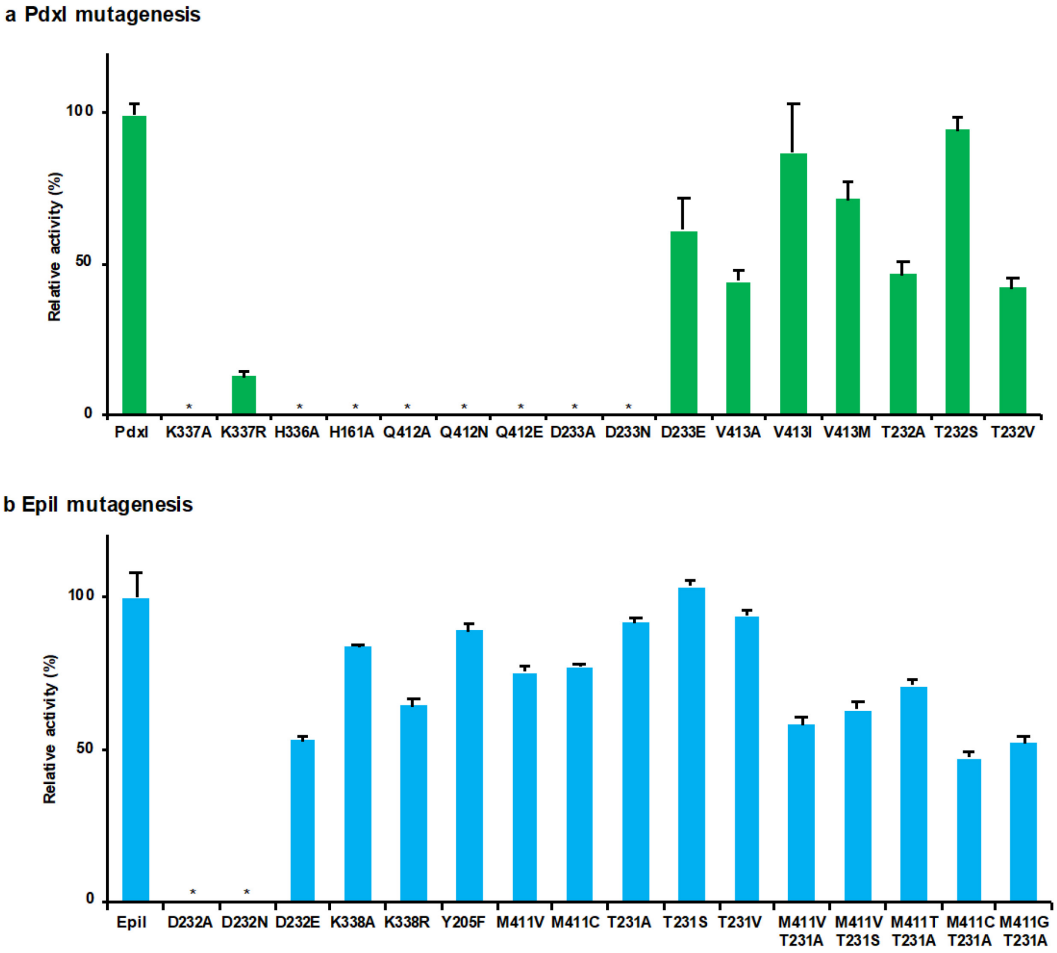
Distances over time of hydrogen bonds to the various positions of the pyridone are tracked in chain A (**a**) and chain B (**b**) of the active site. Left panels show H336 and K337 form hydrogen bonds to the 4-position substituent on the pyridone ring. Right panel shows Q412 and H161 remain hydrogen bonded to 2-position substituent and pyridone nitrogen N1, respectively, for the majority of the simulation. (**c**) Molecular dynamic simulations were initiated from an extended conformation (dihedral = ~180º). Over time, we monitored this conformation to see if the alkyl chain could spontaneously fold to a reactive conformation (dihedral = –20º). Indeed, for short durations of the simulations we observe the chain folding into a reactive conformation for a pericyclic reaction.



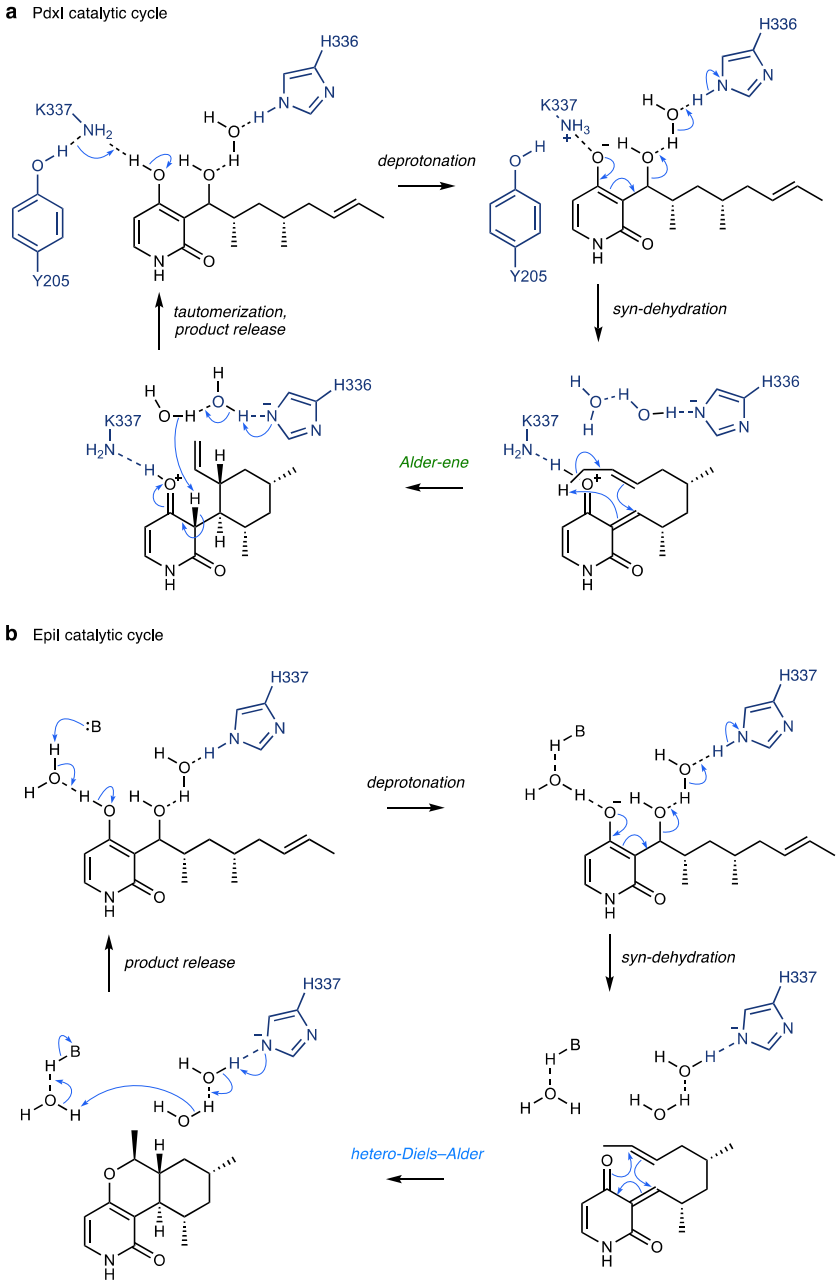
Extended Data Figure 7. HPLC analysis of *in vitro* reaction of PdxI and mutants.
Mutation of the catalytic base K337A abolished the activity, while K337R mutant retained 10% activity. Individual substitution of H336A, Q412A, and H161A all completely abolished the activity. Mutation of D233A or D233N completely abolish enzymatic activity. In contrast, in the D233E mutant 60% of activity and the original periselectivity were retained. This suggests that the carboxylate group of D233 in PdxI is important for enzyme function. A single mutation, V413M is sufficient to change the periselectivity from Alder-ene (>98:2, **8**:**9**) to hetero-Diels–Alder reaction (40:60, **8**:**9**). Further, mutation of T232 to either alanine or serine, but not valine, can lead to the formation of the O2-hetero-Diels–Alder product **11** along with the Alder-ene product **8**. The data show one representative experiment from at least three independent replicates. Reaction conditions: 150 μM **5** with 3 μM PdxG, 1 mM NADPH and 30 μM PdxI (wild type or mutant) at 30 °C for 2 h.



Extended Data Figure 8. HPLC analysis of *in vitro* reaction of EpiI and mutants. In contrast to PdxI, substitution of K338 to alanine did not abolished and retained the activity (83%) (**Extended Data Fig. 9b**). H336A, H161A, and Q410A (corresponding to Q412 in PdxI) mutants were highly insoluble and cannot be assayed. Although D232A and D232N mutations completely abolished the enzymatic activity, the D233E mutation retained 53% of activity and maintained the original periselectivity. This suggests that the carboxylate group of D232 in EpiI is also important for enzyme function. Mutation of Y205F retained 89% activity and maintained the original periselectivity, suggesting the hydroxy group of Y205 is not essential for catalysis. The M411V (corresponding to V413 in PdxI) and M411C mutations increased the Alder-ene product ratio compared to the wild type of EpiI. The T232A and T232S mutations but not T232V mutation, generated the O2-hetero-Diels–Alder product **11** and the Alder-ene product **8** as the minor products, with the hetero-Diels–Alder product **9** as a major product. The double mutation M411V/T231A of EpiI reversed the periselectivity from the native hetero-Diels–Alder reaction (<5:95, **8**:**9**) to the energetically disfavored Alder-ene reaction (2:1, **8**:**9**), although the enzymatic activity is only moderately decreased (**Extended Data Fig. 9**). In the double mutant, **11** was formed due to the mutation of T231. Other double mutants such as M411V/T231S, M411T/T231A, M411C/T231A, and M411G/T231A also reversed periselectivity. The data shown are that of one representative experiment from at least three independent replicates. Reaction Condition: 300 μM **5** with 3 μM PdxG, 1 mM NADPH and 20 μM PdxI mutants at 30 °C for 2 h.



Extended Data Figure 9. **Relative activities of PdxI, EpiI, and mutants.**
The activity of each mutant is compared to that of wild-type PdxI or EpiI quantified by the formation of **8**, **9** and **11**. Error bars indicate s.d. of three independent replicates. Asterisks indicate mutants with no measurable activity. (**a**) The relative enzymatic activity of PdxI mutants. Reaction conditions: 150 μM **5** with 3 μM PdxG, 1 mM NADPH and 30 μM PdxI mutants at 30 °C for 2 h. (**b**) The relative activity of EpiI mutants. Reaction condition: 300 μM **5** with 3 μM PdxG, 1 mM NADPH and 20 μM EpiI mutants at 30 °C for 2 h.



Extended Data Figure 10. Proposed mechanisms of PdxI- and EpiI-catalyzed reactions.
(**a**) The catalytic cycle of PdxI-catalyzed reaction is initiated by the deprotonation of the 4-hydroxy group by K337 followed by the *syn*-dehydration to **7** assisted by the extend water hydrogen bonding network mediated by H336. Subsequently, protonated K337 serves as the hydrogen-bonding catalyst and forms the strong hydrogen bonding with 4-carbonyl oxygen of **7** to set the stage for the periselective Alder-ene reaction. The alkyl chain folds to a reactive conformation and readily undergoes an Alder-ene reaction. After this, the tautomerization is facilitated by K337 and water mediated by H336 to form and release **8**. Then, the next catalytic cycle initiates. (**b**) The catalytic cycle of EpiI-catalyzed reaction, in contrast to PdxI, is initiated by the deprotonation of the hydroxy group by an alternative general base, possibly water followed by the *syn*-dehydration to **7**. Since the key lysine residue does not form hydrogen bonding with 4-carbonyl oxygen of **7**, the favored hetero-Diels–Alder reaction takes place to form and release **9**. Then, the next catalytic cycle initiates.

Methods

Material, fungal strains and culture condition

Aspergillus bombycis NRRL26010 was obtained from Agricultural Research Service Culture Collection (NRRL). *Aspergillus nidulans* FGSC A1145 was obtained from the Fungal Genetics Stock Center (http://www.fgsc.net/). *A. bombycis* and *Epicoccum sorghinum* FT1062 were maintained on PDA (potato dextrose agar, BD) for 3 to 5 days for sporulation or in liquid PDB medium (PDA medium without agar) for isolation of genomic DNA. *A. nidulans* was maintained on Czapek-Dox (CD) agar for sporulation or in liquid CD–ST medium for gene overexpression, compound production and RNA extraction (http://www.fgsc.net/).

General DNA manipulation technique

E. coli TOP10 was used for cloning, following standard recombinant DNA techniques. DNA restriction enzymes were used as recommended by the manufacturer (New England Biolabs, NEB). PCR was performed using Phusion High-Fidelity DNA Polymerase (NEB). The gene-specific primers are listed in Supplementary Information. PCR products were confirmed by DNA sequencing. *E. coli* BL21(DE3) (Novagen) was used as the *E. coli* host for protein expression. GeneArt Seamless Cloning and Assembly kit (Thermo Fisher Scientific) was used for the construction of pet28b-derived protein expression vectors. *In vivo* homologous recombination using *Saccharomyces cerevisiae* was used for the construction of the *A. nidulans* overexpression plasmids.

Protein expression and purification of PdxG, PdxI, EpiI, their homologs and mutants from *E. coli.* BL21 (DE3)

The open reading frame of *pdxI* was amplified using cDNA from the transformant of *A. nidulans* as a template by PCR with primers of pMO20017-f1/r1 and ligated with a linear pet28b expression vector which is amplified using it by PCR with pMOvecf1/r1 to generate pMO20017 using GeneArt Seamless Cloning and Assembly kit (Thermo Fisher Scientific). The identity of the resulting vector pMO20017 was confirmed by DNA sequencing. The codon optimized synthetic gene of *pdxG*, *adxI*, *modxI*, *epiI*, *upiI*, and *hpiI* were synthesized by Integrated DNA Technologies (IDT) co. These synthesized genes were ligated with the linear pet28b expression vector as mentioned above. Construction of plasmids for expression of PdxI and EpiI mutants; The

oligonucleotide primers used for generating the mutants are listed in Supplementary information. The plasmid pMO20017 and pMO90027 containing the wild-type *pdxI* gene and *epiI* gene were used as the template for PCR-based site-directed mutagenesis. The primers of pdx-M-f1/pdx-H161A-r1 and pdx-H161A-f1 /pdx-M-r1 were used for H161A mutant, and the resulting two overlapped fragments and the pET28b expression vector amplified using the primers of pdx-vec-f1/r1 were combined to generate the H161A plasmid using GeneArt Seamless Cloning and Assembly kit (Thermo Fisher Scientific). Other mutants were constructed in the same manner using primer pairs (ex. pdx-M-f1/pdx(or epi)-mutation position-r1 and pdx(or epi)-mutation position-f1/pdx-M-r1). DNA sequencing was used to confirm the identities including the mutated positions of the expression plasmids.

General procedure for overexpression and subsequent protein purification of the proteins of interest was performed as follows: BL21 (DE3) harboring protein expression plasmids was grown overnight in 2 x 5 ml of LB medium with 50 μg/ml kanamycin at 37 °C. Each 2 x 1 L of fresh LB medium with 50 μg/ml kanamycin was inoculated with 5 ml of the overnight culture and incubated at 37 °C until the optical density at 600 nm (OD_{600}) reached 0.6. Then expression of the gene was induced with 100 μM isopropylthio-*β*-D-galactoside (IPTG) at 16 °C. Incubation was continued for another 20 h, after which pellets were collected by using centrifuge and resuspended in lysis buffer (10 mM imidazole, 50 mM Tris-HCl, 300 mM NaCl, pH 8.0) and lysed on ice by sonication. The lysate was centrifuged at 15,000 x g for 30 min at 4 °C to remove the cellular debris. One-step purification of the recombinant C-His_6-tagged fusion proteins from soluble protein by affinity chromatography with Ni-NTA agarose resin (Qiagen) was carried out according to the manufacturer's instructions. Purified proteins were concentrated and exchanged into storage buffer (50 mM Tris-HCl, 300 mM NaCl, 10% glycerol, pH 8.0) with Centriprep filters (Amicon). The purified proteins were checked by SDS-PAGE. Bradford Protein Assay (Bio-Rad) was used to calculate protein concentration.

Protein expression and purification for crystallization

The wild-type PdxI protein was expressed in *E. coli* BL21 (DE3) in LB medium in the presence of 50 μg/mL kanamycin. Expression was induced by 0.4 mM IPTG (isopropyl-*β*-D-thiogalactopyranoside) when OD_{600} reached 1.0. The selenomethionine-derivatized (Se-Met) PdxI was overexpressed in *E. coli* B834 (DE3) in M9 medium at 37 °C in the presence of 50 μg/mL

kanamycin until OD_{600} reached 0.6. After supplementation with 50 mg/L L-(+)-selenomethionine (J&K Scientific), Se-Met protein expression was induced by the addition of 0.4 mM IPTG. After growing for 16 h at 16 °C, the cells were harvested, homogenized in a buffer containing 25 mM tris (pH 8.0), 300 mM NaCl, 5 mM 2-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride (PMSF), and lysed by French press with a high-pressure homogenizer (60-100 MPa). Cell debris was removed by centrifugation at 16000 rpm for 30 min at 4 °C. The supernatant was loaded onto a Ni-NTA affinity column (GE Healthcare), and the His-tagged protein was collected by elution with a buffer containing 25 mM Tris (pH 8.0), 300 mM NaCl, 300 mM imidazole, 5 mM 2-mercaptoethanol, 10% glycerol and further concentrated using Amicon ultra filter units (Millipore) with a 50 kDa molecular weight cutoff. The concentrated protein was then applied to a HiLoad Superdex 200 column (GE Healthcare) in a buffer containing 25 mM tris (pH 8.0), 150 mM NaCl and 3 mM DTT. The peak fractions were collected and concentrated for crystallization and activity assay. The EpiI protein was expressed and purified in the same way as PdxI.

Protein crystallization

Crystals of the Se-Met PdxI were grown at 16 °C using the sitting drop vapor diffusion method in 2 μL drops containing an 1:1 mixture of the protein solution (20 mg/mL Se-Met PdxI in buffer containing 25 mM tris (pH 8.0), 150 mM NaCl and 3 mM DTT) and a reservoir solution (0.1 M MOPS/HEPES-Na (pH 7.5), 12.5% PEG 1000, 12.5% PEG 3350, 12.5% MPD, 0.03 M of each NPS (0.3 M sodium nitrate, 0.3 M disodium hydrogen phosphate, 0.3 M ammonium sulfate)). Stick-like crystals appeared after three days at 16 °C. Crystals of native PdxI were grown at 16 °C using the sitting drop vapor diffusion method in 2 μL drops containing an 1:1 mixture of the protein solution (20 mg/mL PdxI in buffer containing 25 mM tris (pH 8.0), 150 mM NaCl and 3 mM DTT) and a reservoir solution (16% PEG 8000, 40mM potassium phosphate monobasic, 20% glycerol). Stick-like crystals appeared after two days at 16 °C.

All PdxI-compound complexes were crystallized using the sitting drop vapor diffusion method at 16 °C. Purified PdxI (20 mg/mL) was first incubated with 2.0 mM compound and 2-4% DMSO in buffer containing 25 mM tris (pH 8.0), 150 mM NaCl and 3 mM DTT on ice for 30 min and followed by centrifuge. 1 μL of protein solution was mixed with 1 μL of precipitant solution. Stick-like crystals appeared after two days at 16 °C. The PdxI-**5** complex crystals were grown in 0.1 M MOPS/HEPES-Na (pH 7.5), 10% PEG 20 000, 20% PEG MME 550, 0.02 M of each alcohol

(0.2 M 1,6-hexanediol, 0.2 M 1-butanol, 0.2 M (R,S)-1,2-propanediol, 0.2 M 2-propanol, 0.2 M 1,4-butanediol, 0.2 M 1,3-propanediol). The PdxI-**8** complex crystals were grown in the same condition.

The native HpiI crystals were grown in 0.1 M MES (pH 6.5), 15% PEG 6000, 5% MPD. Prism-like crystals appeared after two days at 16 °C. The HpiI-**5** complex was crystallized in 16% PEG 8000, 40 mM potassium phosphate monobasic, 20% glycerol. Prism-like crystals appeared after two days at 16 °C.

All crystals were flash-frozen in liquid nitrogen after being transferred to a cryoprotectant solution consisting of mother liquor supplemented with 10-15% (v/v) glycerol.

Data Collection and Structure Determination

All X-ray diffraction data were recorded at the Shanghai Synchrotron Radiation Facility (SSRF). For Se-Met PdxI, native PdxI, native HpiI and PdxI-**5** complex, data were collected at BL18U1 (λ = 0.97930 Å). For HpiI-**5** complex, data were collected at BL19U1 (λ = 0.97855 Å). For PdxI-**5** complex, data were collected at BL17U1 (λ = 0.97915 Å). Data reduction and integration for PdxI-**8** and HpiI-**5** complex was achieved with XDS[1] while others were achieved with HKL3000 package[2]. The statistics for data collection are listed in Supplementary information. All PdxI crystals belonged to space group $P4_12_12$ with two molecules in the asymmetric unit while all HpiI belonged to space group $C222_1$ with one molecule in the asymmetric unit. Structure of Se Met-PdxI was determined by SAD. 22 selenium sites were identified in one asymmetric unit by Autosol in the PHENIX package[3] and used for model building through Autobuild[3]. Native PdxI structure was determined by molecular replacement using the program PHASER[4] and the atomic coordinates of one chain with the highest completion after Autobuild[3] as the search model. Structures of PdxI complexs, HpiI and HpiI-**5** complex were determined by molecular replacement using PHASER[4] and the atomic coordinates of native PdxI as the search model. Iterative cycles of model rebuilding and refinement were carried out using COOT[5] and PHENIX[3]. PROCHECK[6] and MolProbity[7] were used to access the overall quality of the structural models. Refinement statistics for each final model are recorded in Supplementary Table 1. Structure figures were made using PyMol 2.3 (Schrödinger, LLC)[8].

Activity assay of PdxG-catalysed reaction using the ketone 5 as the substrate

Assays for PdxG activity with 600 μM **5** in HEPES buffer (100 mM HEPES, 2 mM NADPH, pH 8.0) were performed at 50 μL scale with 3 μM PdxG, at 30 °C for 20 min. Then the reaction was quenched with the equal volume of cold acetonitrile. Protein was precipitated and removed by centrifugation and the supernatant analyzed by LC–MS. LC–MS analyses were performed on a Shimadzu 2020 EV LC–MS (Kinetex 1.7 μm C18 100 Å, LC Column 100 × 2.1 mm) using positive- and negative-mode electrospray ionization with a linear gradient of 5–95% acetonitrile MeCN–H_2O with 0.5% formic acid in 15 min followed by 95% MeCN for 3 min with a flow rate of 0.3 ml/min. The result is shown in **Extended Data Fig. 3a**. For the kinetic analysis of PdxG-catalyzed reaction, the 50 μL reaction mixtures containing 3 μM PdxG, 2 mM NADPH, and the different concentrations of **5** (10, 60, 120, 300, 480, 600, 840, 1200 μM) in HEPES buffer were incubated at 30 ºC for 20 min. Then, the reaction was quenched with the equal volume of cold acetonitrile. Protein was precipitated and removed by centrifugation and the supernatant analyzed by LC–MS. The error bars represent standard deviation (s.d.) of three independent replicates. The result is shown in **Extended Data Fig. 3b**.

***In vitro* reaction of PdxI and EpiI using the alcohol 6 as the substrate**

Preparation of the alcohol **6** was performed as follows. To the solution of compound **5** (1.0 mg, 3 μmol) in 0.5 mL EtOH was added $NaBH_4$ (1.2 mg, 30 μmol) at 0 °C and the mixture was stirred at room temperature for 15 min. The reaction mixture was quenched with 0.5 mL 1 M HCl, and the whole was extracted with 0.5 mL ethyl acetate twice. The extract was concentrated, and the residue was subjected to HPLC with an analytical C18 column of Kinetics New column, 5 μm, 4.6 × 250 mm. The fractions were not concentrated due to the instability and immediately used as the substrate solution,

Assays for PdxI and EpiI activity with **6** in HEPES buffer (100 mM HEPES, pH 8.0) were performed at 50 μL scale with 30 μM PdxI (or T232S mutant) or 20 μM EpiI (or T231A mutant), at 30 °C for 2 h. Then the reaction was quenched with 70 μL of cold acetonitrile. Protein was precipitated and removed by centrifugation, and then the supernatant was analyzed by LC–MS as described above. These results are shown in **Extended Data Fig. 4d-e**.

Coupled *in vitro* reaction of PdxG and pericyclases used in this study using 5 as the substrates

In the case of assay for PdxI, the homologs, and the mutants, 50 μL reaction mixture containing 150 μM **5** in HEPES buffer (pH 8.0) containing 100 mM HEPES, 1 mM NADPH was incubated with 3 μM PdxG and 30 μM the pericyclase of interest at 30 °C for 2 h. In the case of assay for EpiI, the homologs, and the mutants, 50 μL reaction mixture containing 300 μM **5** in HEPES buffer (pH 8.0) containing 100 mM HEPES, 1 mM NADPH was incubated with 3 μM PdxG and 20 μM the pericyclase of interest at 30 °C for 2 h. Then, reaction was quenched with 70 μL cold acetonitrile. Protein was precipitated and removed by centrifugation and the supernatant was analyzed by HPLC using a C18 column (Phenomenex Luna C18 (2) 5 μm, 2.0 × 100 mm) with a isoclatic condition of 33% acetonitrile MeCN–H_2O with 0.5% formic acid in 15 min followed by 95% MeCN for 5 min with a flow rate of 0.3 ml/min. The results are shown in **Fig. 1F** and **Extended Data Fig. 2c**. The supernatant was also analyzed by HPLC using a C18 column (Phenomenex Luna C18 (2) 5 μm, 2.0 × 100 mm) with a linear gradient of 15–35% acetonitrile MeCN–H_2O in 4 min followed by 35% MeCN for 11 min followed by 95% MeCN 5 min with a flow rate of 0.3 ml/min. The results are shown in **Extended Data Figs. 7-8**. To measure the relative production ratio of **8**, **9**, and **11** and the relative activity of the pericyclases, the concentrations of **8**, **9**, and **11** were estimated by the each standard curves of **8**, **9**, and **11** that was generated based on peak areas at 290 nm, 290 nm, and 268 nm by HPLC, respectively. The data are shown in **Fig. 3c** and **Extended Data Fig. 9**. The error bars represent standard deviation (s.d.) of three independent replicates.

The measurement of SAM presence in PdxI and EpiI and qualification of the effects of SAM and SAH for PdxI and EpiI catalysis.

To measure whether SAM was copurified with PdxI and EpiI, more than 200 μM of PdxI and EpiI in 50 μL storage buffer (50 mM Tris-HCl, 300 mM NaCl, 10% glycerol, pH8.0) was denatured by acetonitrile. Then, the solutions were centrifuged and the supernatants analyzed by LC–MS. The standards of SAM were also analyzed by LC–MS using a C18 column (Kinetex 1.7 μm C18 100 Å, LC Column 100 × 2.1 mm) with a linear gradient of 5–95% acetonitrile MeCN–H_2O with 0.5% formic acid in 15 min followed by 95% MeCN for 3 min with a flow rate of 0.3 ml/min. The results are shown in **Extended Data Fig. 4a**.

To qualify the effects of SAM and SAH for the PdxI- and EpiI-catalyzed reaction, the 50 μL reaction mixture containing 3 μM PdxG, 1 mM NADPH, and 30 μM PdxI or 20 μM EpiI with or without 1 mM SAM or 1 mM SAH was incubated at room temperature for 10 min. Then, 150 μM **5** was added to the reaction mixture to initiate the enzymatic reaction and the reaction mixtures were further incubated at 30 °C for 2 h. Then, the reaction mixtures were quenched by 70 μL of acetonitrile. After centrifugation, the supernatant was subjected to HPLC analysis using a C18 column (Phenomenex Luna C18 (2) 5 μm, 2.0 × 100 mm) with a isoclatic condition of 33% acetonitrile MeCN–H_2O with 0.5% formic acid in 15 min followed by 95% MeCN for 5 min with a flow rate of 0.3 ml/min. The results are shown in **Extended Data Fig. 4b**.

***In vitro* reaction of PdxI and EpiI using 8 and 9 as the substrates**

To rule out the possibility that PdxI and EpiI catalyze the retro-hydroalkoxylation of **9** to **8** and the hydroalkoxylation of **8** to **9**, respectively, we performed the in vitro reaction of PdxI using **9** as the substrate and EpiI using **8** as the substrate. 50 μL reaction mixture containing 100 μM of the desired compound in HEPES buffer (100 mM HEPES, pH 8.0) was incubated with 30 μM PdxI or 20 μM EpiI at 30 °C for 24 h. Then reaction was quenched with 70 μL acetonitrile. Protein was precipitated and removed by centrifugation and the supernatant analyzed by HPLC using a C18 column (Kinetex 1.7 μm C18 100 Å, LC Column 100 × 2.1 mm) with a linear gradient of 5–95% acetonitrile MeCN–H_2O with 0.5% formic acid in 15 min followed by 95% MeCN for 3 min with a flow rate of 0.3 ml/min. The results are shown in **Extended Data Fig 4c**.

Density functional theory calculations

Initial conformational searches were conducted with Schrodinger's Maestro 2017-2 version 11.2.014.[9] The geometry of conformers were recalculated with the density functional and basis set ωB97X-D/6-31G(d,p) as implemented in Gaussian 16 Rev. A.03 (sse4).[10–16] This functional was chosen for its ability to reproduce CCSD geometry calculations of asynchronous Diels–Alder reactions as well as its general applicability for accurately calculating reaction barriers.[17,18] Structures of interest were further optimized at the reported level of theory, ωB97X-D/6-311+G(d,p).[11,16,19] Following Head-Gordon's suggested basis set for energetics,[18] single point

energies at the ωB97X-D/def2-QZVPP level of theory were computed.[21,22] We believe these methods to accurately calculate energetics for the reported systems and recommend them for use.

Molecular dynamic simulations

Classical molecular dynamics (MD) simulations were performed with the GPU code *pmemd* from the AMBER 16 package.[23] Parameters for ligands were generated within the *antechamber* module with the general Amber force field (gaff) using RESP partial charges calculated at the HF/6-311+G(d,p) level of theory. Each simulation was solvated using the *leap* module in a pre-equilibrated TIP3P truncated octahedral box with a 17 Å buffer and neutralized by Na+ counter ions. Subsequent calculations utilized the Stony Brook modification of the Amber14 force field (ff14sb).[24] All dynamics simulations used a standard minimization, heating, and equilibration protocol before beginning the 500 ns production run.

Data availability

The atomic coordinates of PdxI, PdxI with **5**, PdxI with **8**, HpiI, and HpiI with **5** have been deposited in the Protein Data Bank (http://www.rcsb.org) under the accession code 7BQJ, 7BQK, 7BQL, 7BQP, and 7BQO, respectively.

Methods References

1. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr*. **66**, 125–132 (2010).
2. Minor, W., Cymboriwski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration ofdata reduction and structure solution-from diffraction images to an initial model in minutes. *Acta. Crystallogr. D. Biol. Crystallogr*. **62**, 859–866 (2006).
3. Adams, P. D. et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
4. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Cryst*. **40**, 658–674 (2007).
5. Emsley, P., Lohkamp, B., Scott, W. G & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486−501 (2010).
6. Laskowski, R. A., MacArthur, M.W., Moss, D. S. & Thornton, J. M. ProCheck: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
7. Davis, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**(Web Server issue):W375–W383 (2007).
8. DeLano, W. L. PyMOL: an open-source molecular graphics tool. *Ccp4 Newslett. Protein Crystallogr*. **40**, 11 (2002)
9. Schrödinger Release 2017-2: MacroModel, version 11.2.014; Schrödinger, LLC: New York. (2017).
10. Frisch, M. J. *et al.* Gaussian 16 Revision A.03. (2016).
11. Chai, J.-D. & Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10**, 6615–6620 (2008).

12. Krishnan, R., Binkley, J. S., Seeger, R. & Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **72**, 650–654 (1980).
13. Rassolov, V. A., Pople, J. A., Ratner, M. A. & Windus, T. L. 6-31G* basis set for atoms K through Zn. *J. Chem. Phys.* **109**, 1223–1229 (1998).
14. Francl, M. M. *et al.* Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **77**, 3654–3665 (1982).
15. Dill, J. D. & Pople, J. A. Self-consistent molecular orbital methods. XV. Extended Gaussian-type basis sets for lithium, beryllium, and boron. *J. Chem. Phys.* **62**, 2921–2923 (1975).
16. Hehre, W. J., Ditchfield, R. & Pople, J. A. Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **56**, 2257–2261 (1972).
17. Linder, M. & Brinck, T. On the method-dependence of transition state asynchronicity in Diels–Alder reactions. *Phys. Chem. Chem. Phys.* **15**, 5108–5114 (2013).
18. Mardirossian, N. & Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* **115**, 2315–2372 (2017).
19. Clark, T., Chandrasekhar, J., Spitznagel, G. W. & Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li–F. *J. Comput. Chem.* **4**, 294–301 (1983).
20. Blaudeau, J.-P., McGrath, M. P., Curtiss, L. A. & Radom, L. Extension of Gaussian-2 (G2) theory to molecules containing third-row atoms K and Ca. *J. Chem. Phys.* **107**, 5016–5021 (1997).
21. Schäfer, A., Huber, C. & Ahlrichs, R. Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* **100**, 5829–5835 (1994).
22. Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005).
23. D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman, AMBER 2016, University of California, San Francisco (2016).
24. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).