Structure based engineering for substrate specificity of pinoresinol-lariciresinol reductases

Ying Xiao

Shanghai University of Traditional Chinese Medicine

Kai Shao

Shanghai Institute of Plant Physiology and Ecology

Jingwen Zhou

Jiangnan University <https://orcid.org/0000-0002-3949-3733>

Lian Wang

Jiangnan University

Di Wu

Jiangnan University

Junfeng Chen

Shanghai University of Traditional Chinese Medicine

Jingxian Feng

Shanghai University of Traditional Chinese Medicine

Shi Qiu

Shanghai University of Traditional Chinese Medicine

Lei Zhang

Naval Medical University

Peng Zhang

Shanghai Institute of Plant Physiology and Ecology <https://orcid.org/0000-0003-0408-2923>

Wansheng Chen (✉ chenwansheng@shutcm.edu.cn)

Shanghai University of Traditional Chinese Medicine <https://orcid.org/0000-0002-0025-1315>

Article

Keywords:

Posted Date: July 27th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-41063/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Communications on May 14th, 2021. See the published version at <https://doi.org/10.1038/s41467-021-23095-y>.

Structure based engineering for substrate specificity of pinoresinol-lariciresinol reductases

Running Title: Structural characterization of PLR

Abstract

Pinoresinol–lariciresinol reductases (PLR) are enzymes involved in the lignan biosynthesis after the initial dimerization of two monolignols, which represents the entry point for the synthesis of 8-8' lignans and contributes greatly to their structural diversity. Of particular interest has been in determining how differing substrate specificities are achieved with these enzymes. Here, we present crystal structures of *Li*PLR1/*At*PrR1/*At*PrR2 in the apo, substrate-binding and product-binding forms. Each structure contains a head-to-tail homodimer, and the catalytic pocket comprises structural elements from both monomers. A loop β 4 covers the top of the pocket, and residue 98 from the loop governs catalytic specificity. Structure-guided mutagenesis could switch the substrate specificity of *Li*PLR1 and *At*PrR2, respectively. Our study provides new insight into the molecular mechanism underlying the substrate specificity of PLR/PrRs and suggests an efficient strategy for the large-scale commercial production of the pharmaceutically valuable compound lariciresinol.

Introduction

Lignans are a major group of secondary metabolites in plants¹. This family exhibits numerous biological effects in humans (e.g., anticancer², antiviral³, antioxidant⁴, immunosuppression⁵) owing to their structural diversity—nearly two thousand distinct lignans have been reported. For example, the furofuran lignans such as kandelisesquilignan A/B and terminaloside K have antioxidant effects^{6,7}. The dibenzylbutyrolactone lignans including arctigenin, traxillagenin, arctiin, traxillaside and their glycosides have neuroprotective activities⁸. The aryltetralin lignan podophyllotoxin⁹ is the precursor for the semi-synthesis of anticancer drugs such as Etoposide¹⁰. Lignans biosynthesis starts with the coupling of two coniferyl alcohols by an oxidase (laccase or peroxidase) with the aid of a dirigent protein to form pinoresinol¹¹. Pinoresinol/lariciresinol reductase (PLR), an NADPH-dependent reductase, converts pinoresinol to lariciresinol and subsequently secoisolariciresinol¹². Because the reductive steps that give rise to lariciresinol and secoisolariciresinol represent entry into the biosynthesis of furano, dibenzylbutane, dibenzylbutyrolactone, and aryltetrahydronaphthalene lignan subclasses¹³, PLR is regarded as a pivotal enzyme that contributes to lignan structural diversity. Moreover, variation in both the composition and accumulation of lignans among different plant species, organs and developmental stages can be ascribed, at least in part, to the characteristics of reactions catalyzed by PLRs as well as their expression patterns¹⁴. Therefore, characterization of the catalytic mechanisms of PLRs—especially their substrate selectivities is particularly crucial for understanding the molecular basis of the remarkable diversity of both chemical structures and biological activities of lignans.

The substrate selectivity of PLRs has attracted considerable attention¹². Most PLRs that have been characterized reduce both pinoresinol and lariciresinol efficiently to produce lariciresinol and secoisolariciresinol, respectively¹². Known exceptions are *Arabidopsis thaliana* reductases that have substrate preference toward pinoresinol but only weak (*AtPrR1*) or no activity (*AtPrR2*) toward lariciresinol and are thus named pinoresinol reductases (*PrR*)¹⁵. A recent study indicates the L174I mutant of *Camellia sinensis* PLR1 (*CsPLR1*) loses the capacity to reduce pinoresinol and specifically catalyze lariciresinol, but the underlying mechanism is unclear¹⁶. The three-dimensional structure of *Thuja plicata* PLR1 (*TpPLR1*) has been elucidated and indicates that K138 is responsible for the basic catalysis because the mutant K138A lacks the ability to

convert pinoresinol¹⁷. However, the apo structure does not provide sufficient information to interpret the substrate-selective mechanism of PLR/PrR.

Phylogenetic analysis of PLR/PrR from different species has revealed that *Li*PLR1, a key enzyme involved in lariciresinol biosynthesis in *Isatis indigotica*, has the closest relationship to *At*PrRs with a high level of amino acid identity (>80%)¹⁸ and is grouped with *At*PrR2, which cannot utilize lariciresinol as substrate (Fig. 1). Interestingly, in contrast to *At*PrRs that has substrate preference to pinoresinol, *Li*PLR1 from *I. indigotica* (same family as *A. thaliana*, Cruciferae) can reduce both pinoresinol and lariciresinol efficiently with comparable k_{cat}/K_m values¹⁸. The finding that *Li*PLR1/*At*PrRs, which differ in substrate selectivity, are clustered together suggests that substrate specificity is independent of sequence conservation among PLR/PrR. Therefore, the amino-acid residues responsible for PLR/PrR substrate selectivity are difficult to determine merely through sequence analysis, and thus structural information on PLR/PrR enzymes is vital, as are data concerning how these two enzyme types can utilize two different substrates. In the present work, we compared crystal structures of *Li*PLR1, *At*PrR1 and *At*PrR2 and identified certain residues that may be responsible for the observed substrate selectivity of PLRs and PrRs. Mutagenesis of these residues altered the specificity for pinoresinol and lariciresinol substrates. Our results will enable the synthesis of lignans with diverse chemical structures and bioactivities by biotechnological means or by enzyme-assisted chemistry.

Results

Structural characterization of *Li*PLR1, *At*PrR1 and *At*PrR2

To understand both the catalytic mechanism of PLR and the mechanism underlying the substrate specificity of PLR/PrRs, we chose *Li*PLR1, *At*PrR1 and *At*PrR2 for a structure study. The crystal structures were captured in the apo, substrate-binding and/or product-binding forms (Table 1). We found that, for all 16 structures we solved, each enzyme adopts a similar head-to-tail dimer conformation (Fig. 2a and Supplementary Fig. 1). This arrangement strongly suggests that each PLR/PrR functions as a homodimer, consistent with the literature that *Tp*PLR1 exists as a dimeric entity in solution¹⁷.

Taking the structure of *Li*PLR1-NAP-PIN for description, each protomer contains two domains, namely the N-terminal NADPH binding domain (NBD) and the C-terminal substrate

binding domain (SBD). NBD comprises seven β -strands (β 1–6, β 8) surrounded by six α -helices (α 1–5, α 7), whereas SBD comprises two β -strands (β 7, β 9) with five small α -helices (α 6, α 8–11). A large groove is formed between NBD and SBD (Fig. 2a). Moreover, the groove can be roughly divided into two parts—the positively charged part connecting NBD and the hydrophobic part connecting SBD. The substrates or products can be clearly defined within the groove (Fig. 2a and Supplementary Fig. 2). Several regions (α 5 loop, α 9-helix, α 9 loop) are partially disordered both in the *Li*PLR1_apo and *Li*PLR1_NAP structures, for which the differences can be characterized by a RMSD of 0.287Å, whereas the β 4 loop is well defined in the apo structure (Fig. 2b). Intriguingly, the β 4 loop is well defined in the *At*PrR1_NAP structure but disordered in the *At*PrR1/2_apo structures (Supplementary Fig. 3). These structural differences suggest that the loops are somewhat flexible and act as a switch to control the binding of NADPH and release of NADP⁺. Moreover, the β 2 loop moves slightly towards NADP⁺ in the *Li*PLR1_NAP structure catering to the entry of the coenzyme (Fig. 2b). Residues within the groove form strong hydrogen bonds and hydrophobic interactions with NADP⁺ (Fig. 2c), among which the GXXGXXG motif (considering conserved NADPH binding motif) binds to the phosphate and deoxyribose groups, and residues Ala164, Cys165 together with Phe166 fix the catalytically active nicotinamide group well. The residue Lys144, which corresponds to previously reported Lys138 in *Tp*PLR1 that deduced as a general base for catalysis, forms direct hydrogen bonds with NADP⁺ in *Li*PLR1 or *Li*PLR1_NAP.

Catalytic mechanism of PLR based on its homodimerization

The dimers of *Li*PLR1_NAP and *Li*PLR1_NAP__{+PIN} had similar structures, as suggested by a RMSD of 0.362Å, indeed, even the NADP⁺s could be aligned almost in the same position (Fig. 2d). Both β 4 loops are disordered further implying their flexibility, but β 2 loop and α 10-helix from neighboring molecule of *Li*PLR1_NAP__{+PIN} dimer make contacts with and stabilize the substrate (Fig. 2d-e). Similar conformational changes of β 2 loop and α 10-helix can be seen by comparing the structures for *At*PrR1_NAP__{+PIN} with *At*PrR1_NAP and *At*PrR2_NAP__{+PIN} with *At*PrR2_apo (Supplementary Fig. 4a-b). (+)-Pinoresinol is inserted as a straight chain shape deep into the hydrophobic groove, for which the hydrophilic ends are stabilized through hydrogen bonds formation with main chain atoms of Met125 and Gly178, and the hydrophobic region is surrounded by a series of hydrophobic groups (Fig. 2e). The inner 2-methoxy-phenol group of

(+)-pinoresinol forms sandwich-like π - π stack comprising the nicotinamide head of NADP⁺ and Phe166. Two furan rings in the middle are surrounded by Tyr169 and Phe170 from α 6-helix and His276 and Phe277 form α 10-helix. The outer 2-methoxy-phenol group is coordinated by Phe277 and Val46 of β 2 loop from neighboring protomer, which is distant from the NADP⁺ and (+)-pinoresinol of the protomer (Fig. 2a, 2c, 2e). Further, Lys144 is far away from the furan rings, indicating that it may not participate in catalysis directly. Based on structure analysis, we propose that NADPH probably enters and exits likely through β 4 loop of *Ii*PLR1. Once one molecular (+)-pinoresinol is captured by the narrow hydrophobic groove, each protomer forces the prepositioning of the α 10-helix and β 2 loop of the other protomer, resulting in tight binding of the substrate. This allows H: transfer from the NADPH to the proximal furan ring of the substrate to produce one molecular (+)-lariciresinol.

Regarding to catalytic mechanism of the second step, we further compared the structures of *Ii*PLR1_NAP__{+PIN}, *Ii*PLR1_NAP__{-LAR} and *Ii*PLR1_NAP__{-SEC}, which revealed a similar mode for substrate/product binding (Fig. 2f). Furthermore, Leu46 (corresponding to Val46 in *Ii*PLR1), His276 and Phe277 of *At*PrR1 are positioned similar to the corresponding residues of *Ii*PLR1 to effect substrate binding or product release, except that the β 4 loops cover over the substrate/product, which are disordered in the *Ii*PLR1 structures (Fig. 2d and Supplementary Fig. 4a, 4c-d). The importance of Val46 for catalysis in *Ii*PLR1 is underscored by data from a mutational analysis (Fig. 2g). Mutation of Val46 to Ala resulted in that the conversion of lariciresinol was improved by ~16% and that of secoisolariciresinol was greatly reduced. Mutation V46L of *Ii*PLR1 almost lost the ability to catalyze lariciresinol. These data suggest that *Ii*PLR1 undergoes substrate-induced conformational changes upon homodimerization to achieve catalysis, and the principle of catalytic reactions using lariciresinol as substrate (second step) appears to be like that using pinoresinol as substrate (first step).

Mechanism underlying the substrate selectivity of PLR/PrR

A previous study reported that the recombinant *At*PrR1 could only weakly reduce lariciresinol while that *At*PrR2 lacked activity, which is in sharp contrast to all known PLRs¹⁵. To determine the catalytic mechanism underlying this difference in substrate specificity, we confirmed the relative lack of activity for *At*PrR1/2 (Fig. 1) and then carried out a structure analysis of *Ii*PLR1, *At*PrR1, and *At*PrR2. Each of *Ii*PLR1_NAP__{+PIN}/*At*PrR1_NAP__{+PIN}/*At*PrR2_NAP__{+PIN} is a

homodimer, and superimposition of the protomers among the three complexes revealed RMSDs of 0.374, 0.365, and 0.308Å, respectively (Fig. 3a, 3c and Supplementary Fig. 1). In contrast to *Ii*PLR1_NAP+PIN, the β 4 loops of *At*PrR1_NAP+PIN and *At*PrR2_NAP+PIN can be clearly identified (Fig. 3 and Supplementary Fig. 4). These well-defined loops twist as an “8” shape and cover over both the NADP⁺- and substrate-binding groove. Within the twisted loop, His93 and His97 “grasp” helices α 5 and α 10, while Val92 and Phe94 interact directly with (+)-pinorensinol, and Arg95 strongly interacts with each of NADP⁺, the GXXGXXG motif and α 2-helix from neighboring protomer of the dimer. Similar β 4 loops are also observed in all other *At*PrR1 substrate/product-binding structures, whereas each β 4 loop is disordered in the corresponding *Ii*PLR1 structures, which indicates that the β 4 loop may participate in substrate selectivity and, hence, catalysis. We further explored why the loops behaved differently between *Ii*PLR1 and the *At*PrRs. The amid acid sequences of β 4 loops in the three proteins are quite similar (Supplementary Fig. 5), but residue Ser98 at the C-terminal end of the loop in *Ii*PLR1 is replaced as Asn98 in *At*PrR1/2. Combining sequence and structural data, the difference can be explained reasonably as below: the serine side chain is short enough to remain beneath the guanine group of NADPH, whereas the asparagine side chain cannot do so owing to steric hindrance. Consequently, the asparagine lies nearly vertical to the guanine group and points upward in the structure shown in Fig. 3b and 3d, and thus the swing of β 4 loop is limited in the *At*PrR1_NAP structure. As substrate enters the catalytic site, β 4 loop can fold and cover over the substrate-binding groove (Fig. 3b, 3d and Supplementary Fig. 3b, 4a). Besides the β 4 loop, Val46 in *Ii*PLR1 is replaced with Leu46 in *At*PrR1, which has the effect of compressing the substrate-binding pocket. Although Val46 is unchanged in *At*PrR2, the α 2-helix and β 2 loop from neighbor protomer move closer to the substrate upon its entry into the catalytic site, further condensing the pocket. The relative movement of dimers between *At*PrR2 and *Ii*PLR1 (as suggested by $\sim 9^\circ$ shift shown in Fig. 3c) could be induced by different dimer orientations. Two molecules of the dimer exhibit relative torsion in *At*PrR2, and consequently, Val46 is forced deeper into the substrate-binding pocket compared with what occurs in *Ii*PLR1. Therefore, the entrance and orientation of the substrate in *At*PrR1/2 is more tightly controlled than in *Ii*PLR1.

Mutagenesis-based alteration of the substrate selectivity

Based on the structural analysis of *Ii*PLR1, *At*PrR1 and *At*PrR2, the importance of the candidate

amino acids controlling substrate specificity was verified through site-directed mutagenesis. Enzymatic assays using pinoreosinol as substrate revealed that the *Li*PLR1 mutations including V46A, V46L, S98A and S98H somewhat enhanced the conversion rate of lariciresinol while significantly reduced that of secoisolariciresinol, and mutants V46A, S98A and S98H had > 40% conversion rates for lariciresinol, suggesting that residues 46 and 98 are critical for the substrate preference. Taking V46A as an example for the kinetic analysis, its K_m value for pinoreosinol is 38.9 μM , V_m is 497.8 $\mu\text{M min}^{-1}$, k_{cat} is 17712.6 min^{-1} , and k_{cat}/K_m value is 455.6 $\mu\text{M}^{-1} \text{min}^{-1}$, the latter of which is 500-fold higher than that of wild-type *Li*PLR1¹⁸, indicating that mutant V46A has extremely high catalytic efficiency for pinoreosinol. Consistent with the data for *Li*PLR1, *At*PrR1 mutants L46A and L46V could enhance the conversion rate of lariciresinol and partially reduce that of secoisolariciresinol, which confirmed the importance of these two sites in catalysis. As expected, mutants N98A and N98S in *At*PrR1 had increased activity for secoisolariciresinol production compared with wild type, strongly implying that residue 98 controls the swing of the β 4 loop, which affects substrate binding and catalysis. Interestingly, *At*PrR2 mutant N98S could utilize lariciresinol to produce secoisolariciresinol, with a conversion rate of 1.91%, in contrast to the wild type which lacks this activity. Other *At*PrR2 mutants, including V46A, V46L and N98A, varied in their activities for pinoreosinol, as indicated by the relative rates of conversion to lariciresinol (Fig. 4a). Similar results of were obtained for conversion of lariciresinol to secoisolariciresinol (Fig. 4b). Altogether, the structure-guided mutagenesis indeed could switch the substrate specificity of PLR/PrR, e.g. the *Li*PLR1 mutant V46A had increased preference to pinoreosinol but little catalytic activity for lariciresinol, whereas the *At*PrR2 mutant N98S gained the activity to catalyze lariciresinol (Fig. 4c).

Taking structure and enzymology data together, we proposed a three-step catalytic mechanism for PLR based on its homodimerization. First, the monomeric apo PLR recruits free NADPHs through the very flexible β 4 loop. Second, upon homodimerization, pinoreosinol binds to one protomer via the substrate-binding groove, and the other protomer helps stabilize the substrate. Subsequently, pinoreosinol receives H: from NADPH and be reduced to lariciresinol released later. Third, free lariciresinol is bound by another reductive PLR molecule and fixed by another homodimer, and then the lariciresinol is reduced to secoisolariciresinol and finally released (Fig. 5).

Importantly, the PrRs have more strict requirements for the binding and orientation of lariciresinol compared with PLRs, so PrRs cannot efficiently carry out the third step (Fig. 5). Hence, the substrate-specificity mechanism of PLR/PrRs could be that residues located around the substrate-binding pocket and within the loop, together with residues that promote homodimerization, form the appropriate hydrophobic environment for binding a specific substrate.

Mutational enrichment increases lariciresinol production

Enzymatic assays indicated that certain *Ii*PLR1 mutants could increase the production of lariciresinol from pinoresinol in *E. coli* (Fig. 4). Therefore, these *Ii*PLR1 mutant genes were selected for introduction into pinoresinol-producing *E. coli*¹⁹ for lariciresinol production. Consistent with the enzyme assay results, V46A *Ii*PLR1 produced the greatest amount of lariciresinol (4.23 g/l), which is 18.4-fold greater than that produced by wild-type *Ii*PLR1. Other mutants, including V46L, S98A, S98H and S98N, also achieved a relatively higher yield of lariciresinol compared with *Ii*PLR1 counterpart. These results confirmed that certain *Ii*PLR1 mutants may be good candidates for use in the large-scale production of pharmaceutically valuable lariciresinol. During fermentation, we found that pinoresinol (converted from eugenol) had limited solubility. Therefore, several common surfactants (Tween 20, Tween 80 and Triton X-100) were added individually into the medium to enhance pinoresinol solubility and thereby enhance the subsequent conversion to lariciresinol.

All three surfactants, particularly Tween 80, dramatically increased the production of lariciresinol during *E. coli* fermentation: compared with untreated cells, the addition of Tween 80 increased the yield of lariciresinol from 4.23 to 5.61 g/l in *Ii*PLR1_V46A and from 2.37 to 5.61 g/l in *Ii*PLR1_S98H (Fig. 6). Overall, our results establish a promising route for the production of lariciresinol by synthetic biology strategies.

Discussion

The molecular mechanism of substrate selectivity of PLR/PrR has attracted particular interest owing to the key role of these enzymes in lignan biosynthesis. However, the lack of structural results, especially for PLR/PrR with different substrates, has limited the characterization of the mechanism related to enzyme specificity. In the present study, we characterized crystal structures of *Ii*PLR1, *At*PrR1 and *At*PrR2 in complex with their various substrates. Several residues

participating in substrate binding and catalysis were identified either directly or indirectly based on structural analysis, and these residues were validated by enzyme assays. All these data provide solid evidence to explore the mechanistic basis of substrate selectivity for PLR/PrR. Besides residues 46 and 98 in *Li*PLR1 and *At*PrRs that we identified as being critical for binding and catalysis, residues Phe166, Tyr169, Phe170, His276 and Phe277 within the substrate-binding groove were also strongly correlated with enhanced substrate binding and catalysis. We further deduced that any residue in PLR/PrRs around the hydrophobic groove or affecting homodimerization may impact the conformation of the active site, thereby dictating substrate selectivity (Supplementary Fig. 5-6). Consequently, it is not difficult to understand why mutant L174I of *C. sinensis* PLR1 can hardly reduce pinoresinol and specifically catalyze lariciresinol¹⁶, i.e., because Leu174 points directly toward Tyr163 and thus may indirectly promote substrate recognition.

In addition, PLRs also display substrate stereochemical selectivity, which contributes to enantiomeric diversity of lignans²⁰⁻²³. We found this to be true for *Li*PLR1, which gave comparable k_{cat}/K_m values for both (±)-pinoresinol and (±)-lariciresinol in the range of 0.9–1.6 $\mu\text{M}^{-1} \text{min}^{-1}$, although no experiments with respect to the enantio-specificity of this enzyme were performed^{12,18}. Despite past research on this topic, however, the mechanism underlying the substrate stereochemical selectivity of PLRs remains unclear. Comparison of the enantio-specifically opposite PLRs *Tp*PLR1 and *Tp*PLR2 suggests that F164, V268, and L272 in *Tp*PLR1 contribute to the catalysis of (–)-pinoresinol, whereas L164, G268, and F272 in *Tp*PLR2 prefer to bind (+)-pinoresinol¹⁷. Nevertheless, site-directed mutagenesis carried out in flax indicates that these positions are insufficient to determine enantio-specificity²⁰. Based on amino-acid sequence and structural analyses, it seems that residues Phe94 and Phe277 in *Li*PLR1/*At*PrR1 may act in concert to determine the enantio-specificity of PLRs (Supplementary Fig. 6). These two residues are highly conserved in PLRs that have no enantio-specificity, whereas Ile94 and Tyr277 are present in PLRs that have a strict substrate preference for (+)-pinoresinol. Moreover, there is deletion of two residues on β 4 loop (corresponding to β 4 loop in *Li*PLR1, where Phe94 is located) and leucine in *Lu*PLR1 which has strict enantio-specific for (–)-pinoresinol. Unfortunately, we could not obtain sufficient amounts of the enantiomerically pure substrates to carry out the experiments necessary to establish the enantio-selectivity.

Nature uses a dazzling array of enzymes to produce diverse natural products. However, some modifications are challenging to control because a relative lack of substrate specificity often generates undesired byproducts. *Ii*PLR1 plays an important role in the biotechnological production of lariciresinol¹⁸, which represents the most important component for anti-bacterium, antivirus, and immune regulation effects of the traditional Chinese medicine Radix Isatidis^{3,24,25}. The fact that *Ii*PLR1 can efficiently utilize both pinoresinol and lariciresinol as substrates¹⁸ suggests that the biosynthetic efficiency toward pharmaceutically valuable lariciresinol in Radix Isatidis has been hampered by the relatively low substrate specificity of *Ii*PLR1. In our present work, structure-guided mutagenesis successfully switched the substrate specificity of *Ii*PLR1, leading to overproduction of lariciresinol by *E. coli*. Our study provides new insight into the molecular mechanism underlying the substrate specificity of PLR/PrRs and paves the way for the manufacture of lariciresinol through microbial fermentation. Moreover, this work suggests the possibility of using targeted mutagenesis of *Ii*PLR to improve the efficiency of synthesizing bioactive compounds in *I. indigotica* using current gene-editing technologies^{26,27}.

Methods

Heterologous expression of *Ii*PLR1, *At*PrR1 and *At*PrR2 in *E. coli*

Total RNAs were extracted from leaves of wild-type *I. indigotica* or *A. thaliana* using TRIzol Reagent (GIBCO BRL). The mRNA was reversely transcribed with oligo dT to generate cDNA as a template for PCR. Full-length cDNA sequences of *Ii*PLR1 (GenBank accession no. JF264893), *At*PrR1 (GenBank accession no. AY065214) and *At*PrR2 (GenBank accession no. BT002882) were cloned into pET-duet-1 (Novagen, USA) to generate *Ii*PLR1-pET, *At*PrR1-pET and *At*PrR2-pET, respectively. Purified plasmid DNA was transformed into *E. coli* Rosetta (DE3) cells that grown at 37°C to an OD₆₀₀ of 0.8, after which protein expression was induced by adding 0.5 mM isopropyl β-D-thiogalactoside with incubation overnight at 16°C. Cells were collected, resuspended in buffer A (20 mM Tris-HCl pH 8.0, 100 mM NaCl), and lysed with a French press. The lysate was centrifuged at 20,000 × *g* for 45 min, and the supernatant was applied to a Ni-NTA column equilibrated with buffer A supplemented with 25 mM imidazole. The protein was eluted using buffer A containing 250 mM imidazole and was concentrated for further purification on a Superdex-200 column equilibrated with buffer A. Protein purity was assessed by SDS-PAGE (12%

polyacrylamide), and protein concentration was determined by the Bradford method²⁸.

Crystallization, data collection and structure determination

The full-length *Li*PLR1/*At*PrR1/*At*PrR2 were purified as described above and concentrated to 5-10 mg/ml for crystallization. Aliquots of each concentrated protein sample were mixed 1:1 with reservoir solution, and crystals were grown at 20°C or 4°C in one week using the sitting-drop vapor-diffusion method. For co-crystals, protein was combined with NADP⁺ at a 1:5 molar ratio, and protein with NADP⁺ and substrate/product at a 1:5:10 molar ratio. For reservoir solutions, *Li*PLR1 apo and co-crystals were grown with 0.2M sodium citrate tribasic, 0.1 M sodium citrate/citric acid, pH 4.0 and 20% PEG 3,350; *At*PrR1 apo crystals were grown in 0.2 M lithium chloride, 20% w/v PEG 3,350; *At*PrR1_NAP, *At*PrR1_NAP_₊PIN, *At*PrR1_NAP_₊LAR and *At*PrR1_NAP_₋SEC were grown in 0.2 M sodium fluoride, 20% w/v PEG 3,350; *At*PrR1_NAP_₋PIN and *At*PrR1_NAP_₋LAR crystals were grown in 0.2 M sodium malonate, pH 6.0, 20% w/v PEG 3,350; *At*PrR2 apo crystals were grown in 0.2M magnesium chloride, 0.1 M sodium HEPES, pH 7.5 and 25% PEG 3,350; *At*PrR2_NAP_₊PIN crystals were grown in 2.1 M DL malic acid, pH 7.0. The crystals were cryoprotected by serial transfers into reservoir solutions supplemented with 30% (v/v) glycerol, and flash-cooled in liquid nitrogen. Data collections were performed at BL17U1 and BL19U1 beamline of the Shanghai Synchrotron Radiation Facility. The data were processed with HKL3000²⁹, and the initial phase was determined by molecular replacement with Phenix³⁰, using the crystal structure of *Tp*PLR1 (PDB ID: 1QYD) as a template. The structure models were firstly auto-built in Coot³¹, and were refined by iterative rounds of manual adjustment with Coot and refinement with Phenix. The statistics of data collection and structure refinement are shown in [Supplementary Table 1-4](#).

Site-directed mutagenesis of *Li*PLR1, *At*PrR1 and *At*PrR2 and enzymatic assays

Single-site mutagenesis was achieved through one-step PCR, and mutants were verified with Sanger sequencing. All primers are listed in [Supplementary Table 5](#). After expression and purification of recombinant enzymes under the aforementioned conditions, the results for the enzyme assays for mutants were compared with those for wild-type recombinant *Li*PLR1, *At*PrR1 and *At*PrR2 as follows.

Enzyme activity assays were conducted strictly according to our previous work¹⁸. Assay mixtures (1 ml) consisted of TG buffer (50 mM Tris-HCl, 10% [w/v] glycerol, pH 7.0), 150 μM

NADPH], 100 μ M pinoresinol, or 100 μ M lariciresinol and 5 μ g of purified protein. Assays without fusion protein were used as controls. Protein, buffer, and substrate were pre-incubated for 5 min at 30°C, and each reaction was initiated by addition of NADPH and terminated by addition of 300 μ l ethyl acetate after 30 min. Each assay mixture was extracted with ethyl acetate (3 $\times$ 300 μ l in total). The combined ethyl acetate phases were dried under vacuum, and the residue was dissolved in 1 ml methanol. Conversion rate was then determined. The content of pinoresinol, lariciresinol and secoisolariciresinol was determined by LC-MS using a triple-quadrupole mass spectrometer (Model 6410, Agilent, Santa Clara, CA) following our published methods¹⁸. The selected transitions of m/z were 357→151 for pinoresinol, 359→329 for lariciresinol, and 361→164 for secoisolariciresinol. All the standards were purchased from Sigma-Aldrich (St. Louis, MO).

For determination of $V_{\max}$ and K_m values for *LiPLR1_V46A*, 10 different concentrations of pinoresinol (1–100 μ M) as substrate and 1 μ g purified protein were used. Samples were incubated at 30°C for 5 min (linear kinetic range). Samples without protein were used as controls. The rate of substrate consumption was calculated for kinetic analysis. $V_{\max}$ and K_m values were determined from Lineweaver–Burk plots, and k_{cat} was determined by dividing $V_{\max}$ by the enzyme concentration.

Bioconversion and optimization

Lariciresinol-producing *E. coli* were generated using pinoresinol-producing cells that we had created previously¹⁹. Plasmids with *LiPLR1* mutant genes were transformed into *E. coli* strain D2. The plasmid used was a modified pET-Duet-1, and strain D2 was JM109 carrying plasmid pET28a-PsVAO-Prx02. After culturing on plates, three single colonies identified with colony PCR were picked into 2 ml of LB medium containing ampicillin and kanamycin. These strains were cultured in LB seed culture for 12 h at 37°C and 220 rpm. A 1% seed culture was transferred to a 250-ml shaker flask containing 25 ml TB medium. After culturing for 3 h at 37°C and 220 rpm, 500 μ M isopropyl β -D-thiogalactoside (final concentration) was added to the medium with continued cultivation for 3–3.5 h at 37°C and 220 rpm to OD₆₀₀=0.8. Then 0.75% (v/v) eugenol was added, and a sample was taken after 24 h. We also investigated the effects of several common surfactants on the direct conversion of eugenol to lariciresinol, as follows: eugenol and phosphate-buffered saline (pH 7.0) were mixed 1:1 (v/v), and then 1% (w/v) surfactant (Tween-20,

350 Tween-80 or Triton X-100) was added as an emulsifier, with stirring by an emulsification machine.

351 The concentrations of lariciresinol were determined by HPLC.

352 **Data availability**

353 The atomic coordinates and structure factors for the structures have been deposited in the Protein

354 Data Bank with accession codes AAAA, BBBB, CCCC, DDDD, EEEE, FFFF, GGGG, HHHH,

355 IIII, JJJJ, KKKK, LLLL, MMMM, NNNN, OOOO, PPPP.

356

357 **Competing interests**

358 All the authors have no competing interests.

359

References

1. Teponno, R. B., Kusari, S. & Spiteller, M. Recent advances in research on lignans and neolignans. *Nat. Prod. Rep.* **33**, 1044-1092 (2016).
2. Costa, M. A., Xia, Z. Q., Davin, L. B. & Lewis, N. G. Toward engineering the metabolic pathways of cancer-preventing lignans in cereal grains and other crops. *Recent Advances in Phytochemistry*, (1999).
3. J, L. et al. Lariciresinol-4-O- β -D-glucopyranoside from the root of *Isatis indigotica* inhibits influenza A virus-induced pro-inflammatory response. *J. Ethnopharmacol.* **174**, 379-386 (2015).
4. Xiao, P., Huang, H., Chen, J. & Li, X. In vitro antioxidant and anti-inflammatory activities of Radix Isatidis extract and bioaccessibility of six bioactive compounds after simulated gastro-intestinal digestion. *J. Ethnopharmacol.* **157**, 55-61 (2014).
5. Hirano, T., Wakasugi, A., Oohara, M., Oka, K. & Sashida, Y. Suppression of mitogen-induced proliferation of human peripheral blood lymphocytes by plant lignans. *Planta Med.* **57**, 331-334 (1991).
6. Haihan et al. Two new furofuran lignans from *Kandelia obovata*. *Heterocycles.* **87**, 1093-1098 (2013).
7. Chin, Y., Chai, H., Keller, W. J. & Kinghorn, A. D. Lignans and other constituents of the fruits of *Euterpe oleracea* (Acai) with antioxidant and cytoprotective activities. *J. Agr. Food. Chem.* **56**, 7759-7764 (2008).
8. Jang, Y. P., Kim, S. R. & Kim, Y. C. Neuroprotective dibenzylbutyrolactone lignans of *Torreya nucifera*. *Planta Med.* **67**, 470-472 (2001).
9. Canel, C., Moraes, R. M., Dayan, F. E. & Ferreira, D. Podophyllotoxin. *Phytochemistry.* **54**, 115-120 (2000).
10. Lau, W. & Sattely, E. S. Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science.* **349**, 1224-1228 (2015).
11. Laurence et al. Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. *Science.* **275**, 362-366 (1997).
12. Markulin, L. et al. Pinoresinol-lariciresinol reductases, key to the lignan synthesis in plants. *Planta.* **249**, 1695-1714 (2019).

- 390 13. Dinkova-Kostova, A. T. et al. (+)-Pinoresinol/(+)-lariciresinol reductase from *Forsythia*
391 *intermedia*. Protein purification, cDNA cloning, heterologous expression and comparison to
392 isoflavone reductase. *J. Biol. Chem.* **271**, 29473-29482 (1996).
- 393 14. Umezawa, T. Diversity in lignan biosynthesis. *Phytochem. Rev.* **2**, 371-390 (2003).
- 394 15. Tomoyuki, N., Masaharu, M., Shiro, S., Takefumi, H. & Toshiaki, U. Characterization of
395 *Arabidopsis thaliana* pinoresinol reductase, a new type of enzyme involved in lignan
396 biosynthesis. *J. Biol. Chem.* **283**, 15550-15557 (2008).
- 397 16. Wu, Y. et al. A variable loop involved in the substrate selectivity of pinoresinol/lariciresinol
398 reductase from *Camellia sinensis*. *Phytochemistry*. **162**, 1-9 (2019).
- 399 17. Tongpil, M. et al. Crystal structures of pinoresinol-lariciresinol and phenylcoumaran benzylic
400 ether reductases and their relationship to isoflavone reductases. *J. Biol. Chem.* **278**,
401 50714-50723 (2003).
- 402 18. Xiao, Y. et al. Combined transcriptome and metabolite profiling reveals that *LiPLR1* plays an
403 important role in lariciresinol accumulation in *Isatis indigotica*. *J. Exp. Bot.* **66**, 6259-6271
404 (2015).
- 405 19. Lv, Y., Cheng, X., Du, G., Zhou, J. & Chen, J. Engineering of an H₂O₂ auto-scavenging in
406 vivo cascade for pinoresinol production. *Biotechnol. Bioeng.* **114**, 2066-2074 (2017).
- 407 20. Heimendahl, C. B. I. V. et al. Pinoresinol-lariciresinol reductases with different
408 stereospecificity from *Linum album* and *Linum usitatissimum*. *Phytochemistry*. **66**,
409 1254-1263 (2005).
- 410 21. Chu, A., Dinkova, A., Davin, L. B., Bedgar, D. L. & Lewis, N. G. Stereospecificity of
411 (+)-pinoresinol and (+)-lariciresinol reductases from *Forsythia intermedia*. *J. Biol. Chem.*
412 **268**, 27026-27033 (1993).
- 413 22. Fujita, M., Gang, D. R., Davin, L. B. & Lewis, N. G. Recombinant pinoresinol-lariciresinol
414 reductases from western red cedar (*Thuja plicata*) catalyze opposite enantiospecific
415 conversions. *J. Biol. Chem.* **274**, 618-627 (1999).
- 416 23. Sicilia, T., Niemeyer, H. B., Honig, D. M. & Metzler, M. Identification and stereochemical
417 characterization of lignans in flaxseed and pumpkin seeds. *J. Agr. Food. Chem.* **51**,
418 1181-1188 (2003).
- 419 24. Yang, Z. et al. Antiviral activity of *Isatis indigotica* root-derived clemastanin B against

420 human and avian influenza A and B viruses in vitro. *Int. J. Mol. Med.* **31**, 867-873 (2013).

421 25. Council, N. P. *Pharmacopoeia of China 2015 Edition*. (Chemical Industry, Beijing, 2015).

422 26. Anzalone AV, R. P. D. J. Search-and-replace genome editing without double-strand breaks or
423 donor DNA. *Nature*. **576**, 149-157 (2019).

424 27. Qiupeng Lin, Y. Z. C. X. Prime genome editing in rice and wheat. *Nat. Biotechnol.* **38**,
425 582-585 (2020).

426 28. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of
427 protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254 (1976).

428 29. Cymborowski, M., Otwinowski, Z., Chruszcz, M. & Minor, W. HKL-3000: the integration of
429 data reduction and structure solution--from diffraction images to an initial model in minutes.
430 *Acta Cryst.* **D62**, 859-866 (2006).

431 30. Adams, P. D. et al. PHENIX: a comprehensive Python - based system for macromolecular
432 structure solution. *Acta Cryst.* **D66**, 213-221 (2010).

433 31. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta*
434 *Crystallogr D Biol Crystallogr.* **D60**, 2126-2132 (2004).

435

436 **Table 1.** The crystal structure information of *Ii*PLR1, *At*PrR1 and *At*PrR2 proteins in their apo, NADP⁺ and substrate/product binding forms.
 437

	apo	NADP ⁺	NADP ⁺		NADP ⁺		NADP ⁺	
			(+)-Pinoresinol	(-)-Pinoresinol	(+)-Lariciresinol	(-)-Lariciresinol	(+)-Secoisolariciresinol	(-)-Secoisolariciresinol
<i>Ii</i> PLR1	<i>Ii</i> PLR_apo 2.7 Å	<i>Ii</i> PLR1_NAP 2.4 Å	<i>Ii</i> PLR1_NAP_+PIN 2.3 Å	<i>Ii</i> PLR1_NAP_-PIN 2.1 Å	-	<i>Ii</i> PLR1_NAP_-LAR 2.2 Å	<i>Ii</i> PLR1_NAP_+SEC 2.3 Å	<i>Ii</i> PLR1_NAP_-SEC 2.6 Å
<i>At</i> PrR1	<i>At</i> PrR1_apo 2.8 Å	<i>At</i> PrR1_NAP 2.0 Å	<i>At</i> PrR1_NAP_+PIN 2.0 Å	<i>At</i> PrR1_NAP_-PIN 2.5 Å	<i>At</i> PrR1_NAP_+LAR 1.8 Å	<i>At</i> PrR1_NAP_-LAR 2.5 Å	-	<i>At</i> PrR1_NAP_-SEC 2.0 Å
<i>At</i> PrR2	<i>At</i> PrR2_apo 2.0 Å	-	<i>At</i> PrR2_NAP_+PIN 1.6 Å	-	-	-	-	-

Fig. 1. Biochemical assays of *LiPLR1* and *AtPrRs* function. **a**, *LiPLR1* catalyzes pinoresinol to form lariciresinol and also catalyzes lariciresinol to form secoisolariciresinol efficiently. In contrast, *AtPrR1/2* have substrate preference for pinoresinol but only weak (*AtPrR1*) or no activity (*AtPrR2*) for lariciresinol. **b**, Phylogenetic tree of PLR/PrR from different species. **c**, Conversion of pinoresinol and lariciresinol by recombinant *LiPLR1*, *AtPrR1* and *AtPrR2*. The reaction products were analyzed by LC-MS. **d**, Chromatograms for pinoresinol, lariciresinol, and secoisolariciresinol are denoted in black, red, and blue, respectively.

Fig. 2. Structural mechanism of continuous catalytic reactions by *IiPLR1* based on homodimerization. **a**, Dimer formation of *IiPLR1*_NAP+PIN. Mol-A is shown as cartoon model, with its NBD and SBD colored in light blue and green cyan, respectively. Mol-B is represented as electrostatic-surface model, on which blue and red colors represent positive and negative charges, respectively. NADP⁺ and (+)-pinorexinol are shown as sticks, and colored in orange and yellow, respectively. **b**, Conformational change through comparing monomer structures of *IiPLR1*_apo (grey) and *IiPLR1*_NAP (light orange). The NADP⁺ binding in *IiPLR1*_NAP is colored grey. The $\beta 4$ loops of *IiPLR1*_apo and *IiPLR1*_NAP are highlighted as purple and green, respectively. **c**, Room-in view of NADPH binding groove of *IiPLR1*_NAP. Residues interacting with NADP⁺ are colored as cyan. The conserved GXXGXXG motif is indicated. The Val46 involved in dimer formation and substrate binding is shown as ball-and-stick model and colored as magenta. Possible hydrogen bonds are indicated as dotted lines. **d**, Structure comparison of *IiPLR1*_NAP+PIN and *IiPLR1*_NAP. Mol-Bs of *IiPLR1*_NAP+PIN and *IiPLR1*_NAP, Val46s in *IiPLR1*_NAP+PIN and *IiPLR1*_NAP are colored in marine, orange, magenta, light yellow, respectively. **e**, Room-in view of substrate binding groove. Residues of Mol-As in *IiPLR1*_NAP+PIN and *IiPLR1*_NAP are colored as green cyan (or light blue) and light yellow, respectively. **f**, Structural comparison of substrate/product binding grooves in *IiPLR1*_NAP+PIN, *IiPLR1*_NAP-LAR (blue white), *IiPLR1*_NAP-SEC (smudge). **g**, Enzyme assays for wild-type *IiPLR1* and its mutants.

Fig. 3. Structural differences among *Ii*PLR1, *At*PrR1 and *At*PrR2 indicating different catalytic capacities. **a**, Structural alignment of dimerized *Ii*PLR1_NAP+PIN (green cyan) and *At*PrR1_NAP+PIN (light grey). **b**, Room-in view of dashed box in **a**. **c**, Structural alignment of dimerized *Ii*PLR1_NAP+PIN (green cyan) and *At*PrR2_NAP+PIN (slate). **d**, Room-in view of dashed box in **c**. For brevity, NADP⁺ and some residues in *Ii*PLR1 are hidden.

a

b

c

Fig. 4. Percent conversion of lariciresinol and secoisolariciresinol by *IiPLR1*, *AtPrR1* and *AtPrR2* mutants using pinoresinol or lariciresinol as the substrate. a, Conversion using pinoresinol as substrate. b, Conversion using lariciresinol as substrate. c, LC-MS determination of the products of the reaction catalyzed by *IiPLR1*, *IiPLR1_V46A*, *AtPrR2* and *AtPrR2_N98S*.

Fig. 5. Model depicting the catalytic processes of PLR and PrR.

Fig. 6. Lariciresinol production through introduction of different strains harboring a plasmid encoding *liPLR1* or its single-site mutants into pinoresinol-producing cells.

Figures

Figure 1

Biochemical assays of LiPLR1 and AtPrRs function. a, LiPLR1 catalyzes pinoresinol to form lariciresinol and also catalyzes lariciresinol to form secoisolariciresinol efficiently. In contrast, AtPrR1/2 have substrate preference for pinoresinol but only weak (AtPrR1) or no activity (AtPrR2) for lariciresinol. b, Phylogenetic tree of PLR/PrR from different species. c, Conversion of pinoresinol and lariciresinol by recombinant LiPLR1, AtPrR1 and AtPrR2. The reaction products were analyzed by LC-MS. d, Chromatograms for pinoresinol, lariciresinol, and secoisolariciresinol are denoted in black, red, and blue, respectively

Figure 2

Structural mechanism of continuous catalytic reactions by liPLR1 based on homodimerization. a, Dimer formation of liPLR1_NAP+PIN. Mol-A is shown as cartoon model, with its NBD and SBD colored in light blue and green cyan, respectively. Mol-B is represented as electrostatic-surface model, on which blue and red colors represent positive and negative charges, respectively. NADP⁺ and (+)-pinorelinol are shown as sticks, and colored in orange and yellow, respectively. b, Conformational change through comparing monomer structures of liPLR1_apo (grey) and liPLR1_NAP (light orange). The NADP⁺ binding in liPLR1_NAP is colored grey. The $\beta 4$ loops of liPLR1_apo and liPLR1_NAP are highlighted as purple and green, respectively. c, Room-in view of NADPH binding groove of liPLR1_NAP. Residues interacting with NADP⁺ are colored as cyan. The conserved GXXGXXG motif is indicated. The Val46 involved in dimer formation and substrate binding is shown as ball-and-stick model and colored as magenta. Possible hydrogen bonds are indicated as dotted lines. d, Structure comparison of liPLR1_NAP+PIN and liPLR1_NAP. Mol-Bs of liPLR1_NAP+PIN and liPLR1_NAP, Val46s in liPLR1_NAP+PIN and liPLR1_NAP

are colored in marine, orange, magenta, light yellow, respectively. e, Room-in view of substrate binding groove. Residues of Mol-As in liPLR1_NAP_+PIN and liPLR1_NAP are colored as green cyan (or light blue) and light yellow, respectively. f, Structural comparison of substrate/product binding grooves in liPLR1_NAP_+PIN, liPLR1_NAP_-LAR (blue white), liPLR1_NAP_-SEC (smudge). g, Enzyme assays for wild-type liPLR1 and its mutants.

Figure 3

Structural differences among liPLR1, AtPrR1 and AtPrR2 indicating different catalytic capacities. a, Structural alignment of dimerized liPLR1_NAP_+PIN (green cyan) and AtPrR1_NAP_+PIN (light grey). b, Room-in view of dashed box in a. c, Structural alignment of dimerized liPLR1_NAP_+PIN (green cyan) and AtPrR2_NAP_+PIN (slate). d, Room-in view of dashed box in c. For brevity, NADP+ and some residues in liPLR1 are hidden.

Figure 4

Percent conversion of lariciresinol and secoisolariciresinol by liPLR1, AtPrR1 and AtPrR2 mutants using pinoresinol or lariciresinol as the substrate. a, Conversion using pinoresinol as substrate. b. Conversion using lariciresinol as substrate. c, LC-MS determination of the products of the reaction catalyzed by liPLR1, liPLR1_V46A, AtPrR2 and AtPrR2_N98S.

Figure 5

Model depicting the catalytic processes of PLR and PrR.

Figure 6

Lariciresinol production through introduction of different strains harboring a plasmid encoding liPLR1 or its single-site mutants into pinioresinol-producing cells.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarydata.pdf](#)
- [reportingsummaryNCOMMS2027597.pdf](#)
- [PDBvalidationreports.zip](#)