Controlling Chemo- and Regio-selectivity of a plant P450 in yeast cell towards rare licorice triterpenoids biosynthesis

Sun, W., Xue, H., Liu, H., Lv, B., Yu, Y., Wang, Y., Huang, M., & Li, C. (2020). Controlling Chemo- and Regio-selectivity of a plant P450 in yeast cell towards rare licorice triterpenoids biosynthesis. *ACS Catalysis*. https://doi.org/10.1021/acscatal.0c00128

Published in:
ACS Catalysis

Document Version:
Peer reviewed version

Controlling Chemo- and Regio-selectivity of a plant P450 in yeast cell towards rare licorice triterpenoids biosynthesis

Journal:	*ACS Catalysis*
Manuscript ID	cs-2020-00128b.R2
Manuscript Type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Sun, Wentao; Beijing Institute of Technology Xue, Haijie; Beijing Institute of Technology liu, Hu; Beijing Institute of Technology Lv, Bo; Beijing Institute of Technology, Department of Biological Engineering Yu, Yang; Beijing Institute of Technology Wang, Ying; Beijing Institute of Technology, Department of Biological Engineering Huang, Meilan; Queen's University Belfast, School of Chemistry and Chemical Engineering Li, Chun; Beijing Institute of Technology; Tsinghua University

Controlling Chemo- and Regio-selectivity of a plant P450 in yeast cell towards rare licorice triterpenoids biosynthesis

Wentao Sun,[†] Haijie Xue,[†] Hu Liu,[†] Bo Lv,[†] Yang Yu,[*,†] Ying Wang,[*,†] Meilan Huang,[§] and Chun Li[*,†,‡]

†. Institute for Synthetic Biosystem/Department of Biochemical Engineering, School of Chemistry and Chemical Engineering, Beijing Institute of Technology, Beijing 100081, P. R China

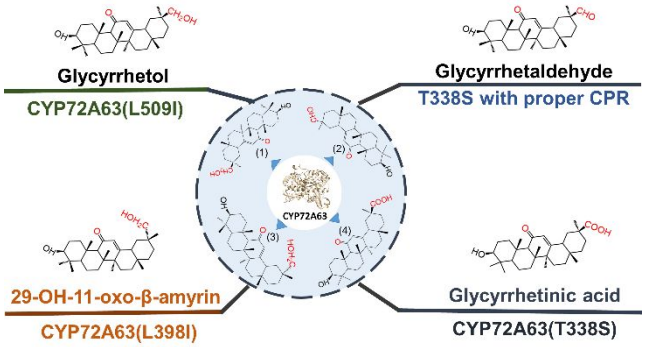
‡. Key Lab of Industrial Biocatalysis Ministry of Education, Department of Chemical Engineering, Tsinghua University, Haidian District, Beijing, 100084, P. R. China

§. School of Chemistry and Chemical Engineering, Queen's University Belfast, Stranmillis Road, Belfast, BT9 5AG, Northern Ireland, United Kingdom

*Corresponding author: lichun@tsinghua.edu.cn (Chun Li)

*Co-corresponding author: wy2015@bit.edu.cn, yangyu1@bit.edu.cn

ABSTRACT: The hallmark reaction of P450 monooxygenase involves activation of C-H bond and the production of a hydroxyl. P450s tailoring natural product could further oxidize the hydroxyl to carboxylic acid. However, heterogeneously expressed plant P450s display poor chemo- and regio-selectivity in microbes, restricting the efficient biosynthesis of related natural products. CYP72A63 is a P450 enzyme previously used for the biosynthesis of glycyrrhetinic acid, and its poor selectivity resulted in oxidation of 11-oxo-β-amyrin to a mixture of rare licorice triterpenoids (glycyrrhetol, glycyrrhetaldehyde, glycyrrhetinic acid and 29-OH-11-oxo-β-amyrin). In this study, we have identified key residues, which influence the enzyme-substrate hydrophobic interaction, in controlling the chemo- and regio-selectivity of the enzyme, and engineered the enzyme toward selectivity oxidation to hydroxyl and carboxylic acid. Moreover, tuning the redox partner of the P450 leads to selective production of glycyrrhetaldehyde, a good starting point for further modification. In this study, controlling the catalytic property of plant P450s prove to be of great use in the synthesis of desired licorice triterpenoids, which can be used in biosynthesis of other terpenoid natural products.



(Abstract graph)

Key words: cytochrome P450; rare licorice triterpenoids; protein engineering; enzyme selectivity; yeast

INTRODUCTION

Cytochrome P450s are capable of activating C-H bonds and oxidize it to hydroxyl or carboxyl group with high chemo- and regio-selectivity.[1] They are important tailoring enzymes in natural product synthesis[2]. Although the reaction mechanism is well-studied using model bacterial P450s,[3-10] the underlying molecular basis for chemo- and regio-selectivity remain elusive, especially for plant P450s,[11] which are challenging to be recombinantly expressed and purified.[12]

Glycyrrhizin, a triterpenoid from licorice roots, is widely used as anti-virus,[13-14] hepatoprotective,[15] anti-allergic[16] and antiulcer[17] drugs and could be produced in large scale through phytoextraction.[18-20] Compared with glycyrrhizin, the rare licorice triterpenoids including the precursors of glycyrrhizin such as glycyrrhetinic acid, glycyrrhetaldehyde, glycyrrhetol (30-OH-11-oxo-β-amyrin) and related isomers have extremely low abundance in licorice, and possess more important physiological activities.[21] Additionally, the functional groups of -CHO, -OH on the skeleton can be chemically modified, expanding the structural and functional diversities of glycyrrhetinic acid-like molecules.[22] Nevertheless, these compounds are hard to obtain by either phytoextraction or chemical synthesis, for their extremely low concentration in plant and complex structures.

We have employed a P450 monooxygenase from *Medicago truncatula*, CYP72A63, to produce glycyrrhetinic acid in yeast. The enzyme could produce a mixture of glycyrrhetinic acid, glycyrrhetol, glycyrrhetaldehyde, as well as the glycyrrhetol isomer 29-OH-11-oxo-β-amyrin[23] in the host cell.[24-26] Their structural similarity makes it hard to separate and purify these rare licorice triterpenoids.

In this study, we combined computational and experimental approaches to improve the *in vivo* activity, regio- and chemo-selectivity of CYP72A63. Key residues were identified by homology modeling and molecular docking. Mutations on these residues, along with matching redox partner of P450, convert the promiscuous P450 enzyme into an enzyme oxidizing a specific C-H bond into hydroxyl, aldehyde and carboxylic acid, respectively. These results provide promising solutions to produce rare licorice triterpenoids and other derivatives, and provide strategies on fine-tuning the reactivity of other plant P450s to synthesize target compounds in yeast.

RESULTS AND DISCUSSION

Controllable regioselectivity and consecutive oxidative ability for specific synthesis of glycyrrhetinic acid

Although CYP72A63 was the best enzyme for the oxidation of 11-oxo-β-amyrin to produce glycyrrhetinic acid reported so far, the low selectivity limits its efficiency.[24] The yeast strain GA0 (Table S1) was constructed by introducing *CYP72A63* into a host (SynV) together with genes on the 11-oxo-β-amyrin synthesis pathway. Glycyrrhetol, glycyrrhetaldehyde, glycyrrhetinic acid and

the isomer 29-OH-11-oxo-β-amyrin were simultaneously produced, with a product ratio of 45.2: 21.4: 29.0: 4.40 (**Figure 1A, Figure S1, S2)**. The titer of the end oxidation product, glycyrrhetinic acid was 4.2±0.4 mg/L. According to the product distribution, the rate-limiting step in the consecutive oxidation of C-H bond at C-30 is the further oxidation of hydroxyl. A small fraction of substrate was oxidized at the C-29 rather than at the C-30, further diverted the substrate to an undesired direction and complicated the downstream purification. On the other hand, the catalytic promiscuity of CYP72A63 provides a starting point for rational design of the enzyme toward different activities.

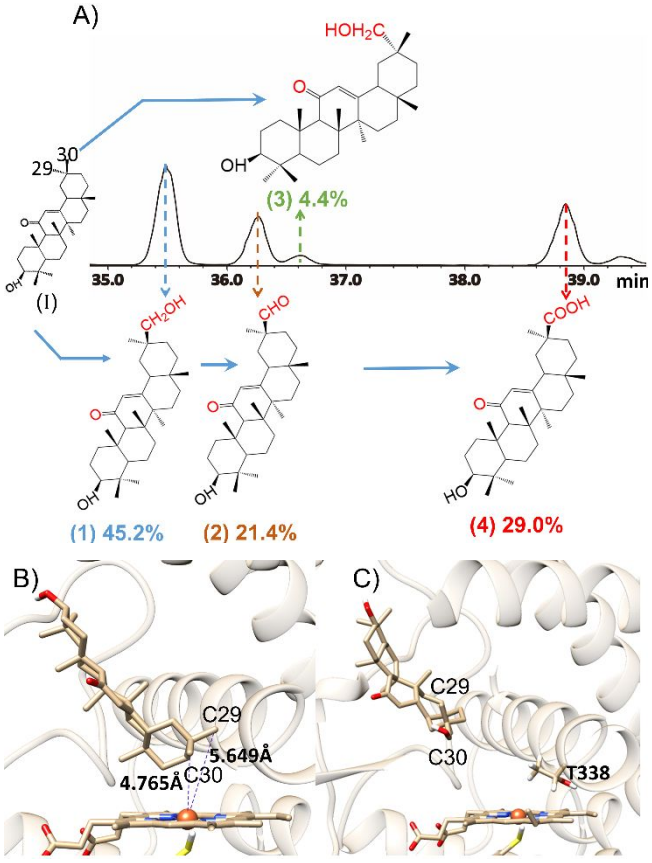


Figure 1 (A) GC-MS spectrum showing the product distribution in yeast strain harboring wild type CYP72A63.(**I**) 11-oxo-β-amyrin, the substrate of CYP72A63, (1) glycyrrhetol, (2) glycyrrhetaldehyde, (3) 29-OH-11-oxo-β-amyrin, (4) glycyrrhetinic acid and the number in % indicates the percentage of the specific oxidized product in the total oxidized products from 11-oxo-β-amyrin. (B) Structure model showing 11-oxo-β-amyrin in wild type CYP72A63, (C) Structure model showing glycyrrhetol in the wild type CYP72A63, indicating that the C-30 hydroxyl of glycyrrhetol was unfavorably located due to the hydrophobic force of the proximal active pocket with threonine (T338) methyl orienting to the substrate.

Recombinant expression and purification of CYP72A63 have been attempted, but the protein failed to be expressed in *Escherichia coli.* Due to the difficulty in obtaining the purified enzyme and the corresponding substrates, we did not carry out *in vitro* activity assay. Alternatively, we expressed the enzyme and its mutants in the engineered yeast, in which the substrate, 11-oxo-β-amyrin, accumulates to a high level (>70 mg/L) with engineered metabolic pathways consisting of bAS (β-amyrin synthase), uni25647, and GuCPR1 (GA0, Table S1).[24]

As an alternative to experimentally determining the structure, homology models of CYP72A63 were created to further understand the determining factors of its catalytic characteristics, and substrates were subsequently docked into the model by Chimera 1.13.[27] The structural model was evaluated by PROCHECK[28] and ProSA[29-30]. PROCHECK showed that the model is of excellent quality with 90.4% of the residues in the most favored region of the Ramachandran plot. The ProSA energy graph is in negative range and calculated z-sorce (-8.4) was in the range of natively folded protein.

The structure model with 11-oxo-β-amyrin docked into CYP72A63 showed that the distances of the C-30 and C-29 to heme-Fe were 4.765 Å and 5.649 Å, respectively **(Figure 1B)**, both of which were well within the range accessible for other P450s.[31-32] Hence, both of the two moieties were expected to be hydroxylated and the C-30 was preferred for its proximity to the heme-Fe. The model was in accordance with the experimental result that both C-29 and C-30 were oxidized, and the C-30 oxidation is preferred.

To further investigate the consecutive oxidation of 11-oxo-β-amyrin to glycyrrhetol, glycyrrhetaldehyde, and glycyrrhetinic acid, we docked glycyrrhetol with CYP72A63 **(Figure 1C)**., Glycyrrhetol's hydrophilic C-30 moiety with a hydroxyl was pushed away from the heme-Fe due to the hydrophobic force of the proximal active pocket with threonine methyl orienting to the substrate **(Figure 1C)**. Compared with 11-oxo-β-amyrin, the glycyrrhetol was less fit in CYP72A63. As a result, glycyrrhetol should be harder to be oxidized than 11-oxo-β-amyrin. It is consistent with the experimental result that the glycyrrhetol is accumulated over glycyrrhetinic acid.

Since 11-oxo-β-amyrin is a bulky and rigid molecule, it could not be freely repositioned in the active pocket to alter the regioselectivity between adjacent moieties like a smaller molecule. Considering the hydrophobicity difference between 11-oxo-β-amyrin and glycyrrhetol, the binding

pocket of CYP72A63 needs to be tailored to make it less hydrophobic to push C-29 away from the heme-Fe for the selective oxidation on C-30 while pulling the glycyrrhetol closer to heme-Fe to promote further oxidation.

From the structure model of CYP72A63 docked with substrates, we identified L330, L333, T338 and A334 as residues in the binding pocket potentially influencing the hydrophobic interaction between binding pocket and substrate **(Figure S3)**. Mutagenesis of these residues (L330A/T/S/E, L333 A/T/S/E, A334 T/S/E and T338A/S) revealed T338's key role in the determination of catalytic property of the enzyme. Due to its proposed role in proton transfer (*vide infra*), the T338A mutant almost abolished the activity **(Figure S4)**. Mutating threonine to serine keeps its proton delivery activity, while decreases the hydrophobicity slightly.[33] Indeed, experimental result showed a yield of 30.0±0.4 mg/L glycyrrhetinic acid in the engineered yeast harboring CYP72A63 (T338S), which was 7.1-folds higher than that in the control strain using the wild type CYP72A63. More importantly, the strain with CYP72A63 (T338S) produces 5.8% glycyrrhetaldehyde and 94.2% glycyrrhetinic acid **(Figure 2C)**, showing full regioselectivity toward C-30 oxidation and more complete oxidation. To avoid the potential endogenic factor of yeast on the catalytic property of the enzyme, time-course experiments was conducted. Compared with the wild type CYP72A63, the mutant led to more completed oxidation and full regioselectivity in the time-course experiments, further indicating the significant improvement of the catalytic property. (Figure S5)

To further investigate the mechanism of the better activity of T338S mutant, the homology model was created and substrates (11-oxo-β-amyrin and glycyrrhetol) were docked into the model by Chimera 1.13, respectively. The results showed that the serine hydroxyl points to the pocket, which is more hydrophilic than the threonine methyl group for the wild type enzyme **(Figure 2 A, B)**. The change on hydrophobicity expanded the repulsive interaction between 11-oxo-β-amyrin and the proximal active pocket close to heme and I-helix. As a result, the distance between the C-30 and C-29 to the heme-Fe increased to 5.217 Å and 6.681 Å, respectively **(Figure 2 A)** while the angle of C-30, C-29 and the heme-Fe decreased from 22.4° to 19.6°. Therefore, the C-29 oxidation was avoided by mutating T338 to serine. Compared with the wild type CYP72A63 **(Figure 1C)**, the C-30 hydroxyl of glycyrrhetol in the mutant CYP72A63 (T338S) was pulled closer to the heme-Fe with a proper orientation **(Figure 2B)**, so that alcohol intermediates could be oxidized furtherly to

obtain glycyrrhetinic acid resulting in an efficient consecutive reaction. This approach provided a model to control the regioselectivity and activity of CYP450s for the thorough C-H oxidation of relatively large substrate.

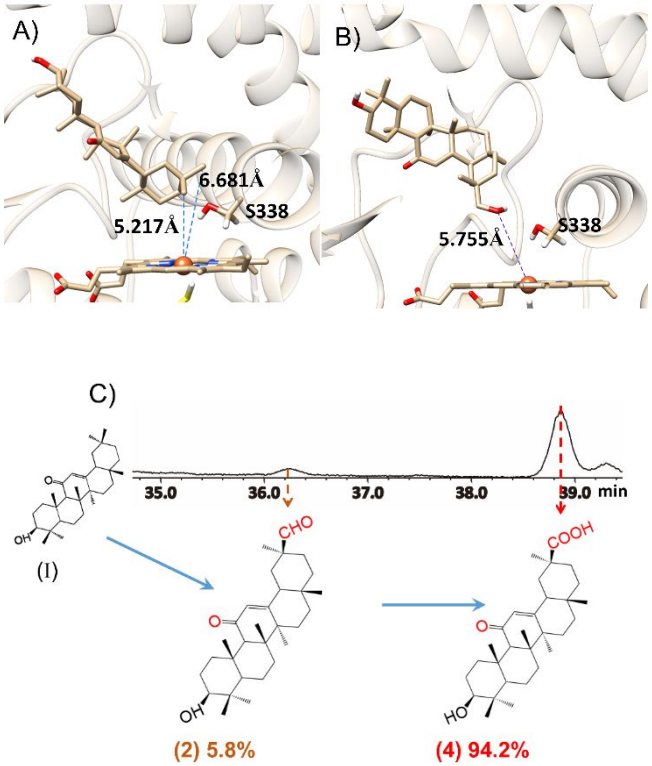


Figure 2. (A) The structure model of CYP72A63(T338S) with 11-oxo-β-amyrin, (B) The structure model of CYP72A63 (T338S) with glycyrrhetol, (C) GC-MS spectrum showing the product distribution in yeast strain harboring CYP72A63 (T338S), (**I**) 11-oxo-β-amyrin, the substrate of CYP72A63 (1) glycyrrhetol, (2) glycyrrhetaldehyde (4) glycyrrhetinic acid, and the number in % indicates the percentage of the specific oxidized product in the total oxidized products from 11-oxo-β-amyrin.

Controllable chemoselectivity for specific production of glycyrrhetol

Glycyrrhetol has excellent anti-virus activity but lower cytotoxicity than glycyrrhetinic acid,[34] and the hydroxyl on C-30 would expand the chemical space for further modification to produce novel bioactive derivatives. To realize the biosynthesize of glycyrrhetol, the oxidation of 11-oxo-β-amyrin should be terminated at the alcohol intermediate, which means to decrease the affinity of CYP72A63 to glycyrrhetol while keeping the affinity to 11-oxo-β-amyrin. Compared with glycyrrhetol, 11-oxo-β-amyrin has a better orientation in the proximal active pocket (**Figure 1**), indicating a natural preference of 11-oxo-β-amyrin than glycyrrhetol. Thus, we kept the proximal

active pocket intact during engineering. Besides the residues lining in the active pocket, residues distant from the active pocket could play an important role in the substrates recognition and binding.[35-36] We hyposized that increasing the hydrophobicity of the distal term of active pocket would disturb the substrate recognition, which would further inhibit the binding of glycyrrhetol.

Based on the structure model, we mutated L509, L137, and L242 located above the distal active pocket to isoleucine and valine with increased hydrophobicity **(Figure 3A)**. CYP72A63 (L242I) and CYP72A63 (L509I) mutants display an enhanced chemo-selectivity compared with the wild type **(Table 1)**, with CYP72A63 (L509I) mutant producing 90.6% of glycyrrhetol among the oxidation products of 11-oxo-β-amyrin **(Figure 3B)**. We then mutated the 509th residue to all other amino acids and tested all the mutants in the yeast cell. Among the obtained mutants, L509T/H/N with decreased hydrophobicity also showed sharply increased glycyrrhetol ratio **(Table 2, Figure S6).** However, the titers of glycyrrhetol in yeast strains with CYP72A63 (L509T/H/N) are much lower. The underlying reason is probably that the more hydrophilic sidechain disrupted the binding of the 11-oxo-β-amyrin resulting in a sharply decreased total oxidation ability. These results suggest that adjusting the substrate accessibility is an efficient approach to control the chemo-selectivity of P450 for the desired oxidation intermediates.

Table 1. Products spectrum of mutants on L509 and L242

Mutants	glycyrrhetol %*	glycyrrhetaldehyde %*	29-OH-11-oxo-β-amyrin %*	glycyrrhetol titer mg/L
L509I	90.6	4.9	4.5	6.3±1.2
L509T	87.8	6.2	6.0	0.96±0.12
L509H	100	0	0	0.48±0.06
L509N	100	0	0	0.58±0.03
L242I	76.6	14.6	8.8	5.7±0.95

* The % means the percentage of the total oxidized products of 11-oxo-β-amyrin

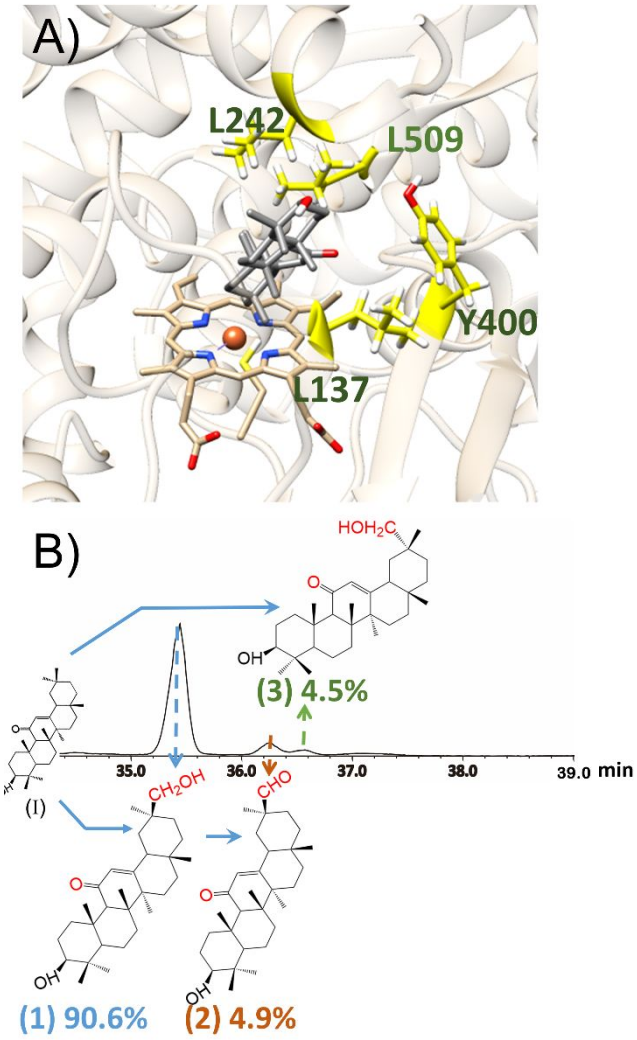


Figure 3. (A) L509 locates in the distal active pocket, (B) GC-MS spectrum showing the product distribution in yeast strain harboring CYP72A63 (L509I). (**I**) 11-oxo-β-amyrin, the substrate of CYP72A63 (1) glycyrrhetol, (2) glycyrrhetaldehyde, (3) 29-OH-11-oxo-β-amyrin and the % indicates the percentage of the total oxidized products of 11-oxo-β-amyrin

CPR matching for specific synthesis of glycyrrhetaldehyde

Aldehyde is an intermediate for the oxidation of C-H bond to carboxylic acid. With high reactivity, the aldehyde group provides a handle for convenient derivatization and has become an important functional group for bioorthogonal chemistry[37]. Due to its high reactivity, there is rare P450 reported so far to accumulate aldehyde from a C-H bond as the main product. To stop further oxidation of glycyrrhetaldehyde, the accessibility of the glycyrrhetaldehyde to CYP72A63 should be reduced while keeping the glycyrrhetol as a preferred substrate. Since glycyrrhetaldehyde and

glycyrrhetol are similar in hydrophobicity and other properties, direct engineering of P450, such as reshaping the active pocket and mutating residues in the distal active pocket did not give satisfactory result.

Cytochrome P450 reductases (CPR) modulate the activities of P450s through influencing the electron transfer and substrate binding process.[38] Moreover, CPRs can affect the catalytic properties of P450s since their interaction might lead to the conformational change of the P450 as well as influence the reaction process.[39-40] In our previous research, several CPRs from different leguminous plants were screened in order to boost the activity of CYP72A63.[24] Some of them led to low glycyrrhetinic acid production but high alcohol intermediates synthesis, which were called as low efficiency CPRs. We tested four low efficiency CPRs, i.e. AtCPR1 from *Arabidopsis thaliana*, MtCPR2 and MtCPR3 from *Medicago truncatula*, and GuCPR2 from *Glycyrrhiza uralensis*, with the mutant CYP72A63 (T338S), respectively, and then analyzed the aldehyde production in yeast. Increased ratios of aldehyde were observed in CYP72A63 (T338S) with low efficiency CPRs including GuCPR2, AtCPR1, MtCPR2 and MtCPR3 **(Table 2)**, indicating a significant influence of CPRs on the chemo-selectivity of the CYP72A63 (T338S). The CYP72A63 (T338S) mutant with GuCPR2 showed the best selectivity with 82.3% of oxidation production being aldehyde and an aldehyde yield of 31.4±5.8 mg/L **(Figure 4)**. Moreover, for instant derivatization of glycyrrhetaldehyde, the crude extract of the cultured cell was incubated with 2,4-Dinitrophenylhydrazine at pH5~6 for 30 min. The reaction mixture was tested by HPLC with a DAD detector. New peaks with the absorption maximum around the 400nm appeared **(Figure S7)**, which is the characteristic absorption of the hydrazone product formed from condensation of hydrazine and aldehyde. The reaction suggests the aldehyde provides an excellent handle for convenient derivatization of terpenoid compounds. To our knowledge, our work represents the first P450 enzyme producing aldehyde with high selectivity. The production of aldehyde in yeast by pairing various CPRs with P450 in our study provides an alternative strategy to fine-tune specific P450s for different target intermediates and further derivatization.

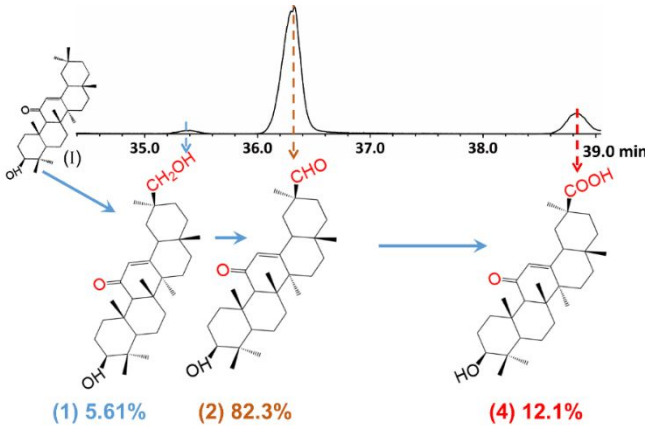


Figure 4. GC-MS spectrum showing the product distribution in yeast strain harboring CYP72A63 (T338S) and GuCPR2: (**I**) 11-oxo-β-amyrin, the substrate of CYP72A63 (1) glycyrrhetol, (2) glycyrrhetaldehyde (4) glycyrrhetinic acid, and the % means the percentage of the total oxidized products of 11-oxo-β-amyrin.

Table 2. **The products of yeast whole-cell catalyst carrying CYP72A63 (T338S) and different CPRs**

CPR	glycyrrhetol %*	glycyrrhetaldehyde %*	glycyrrhetinic acid %*
AtCPR1	2.2	70.3	27.5
MtCPR2	12.1	77.5	10.4
MtCPR3	1.4	71.8	26.8
GuCPR2	5.6	82.3	12.1

* The % means the percentage of the total oxidized products of 11-oxo-β-amyrin

Rotation of substrate orientation for specific synthesis of 29-OH-11-oxo-β-amyrin

29-OH-11-oxo-β-amyrin is isomeric to glycyrrhetol. The change of position of hydroxyl group might lead to novel bioactivity. As both glycyrrhetol and 29-OH-11-oxo-β-amyrin are produced by the wild type CYP72A63, we investigated key factors in determining the regioselectivity of the enzyme. The structure model of the wild type CYP72A63 showed that the C-30 moiety was closer to the heme and easier to be oxidized. The regioselectivity could be reversed to specifically oxidize C-29 if 11-oxo-β-amyrin is rotated in the active pocket. L398 and L149 flanking the E-ring were identified as possible residues to be useful to push the E-ring of the 11-oxo-β-amyrin and thus rotate it in the active pocket **(Figure 5A)**. To test our hypothesis, site-specific saturation mutagenesis of L398 and L149 were carried out and all the mutants were introduced into the engineered yeast.

GA170 harboring mutant CYP72A63 (L398I) specifically produced 10.2±2.2 mg/L of 29-hydroxyl-11-oxo-β-amyrin with 100% regioselectivity **(Figure 5C)**. To further investigate the molecular basis of the improved regioselectivity, 11-oxo-β-amyrin was docked into the homologous model of CYP72A63 (L398I). The docking results showed that the I398 led to a sharp rotation of 11-oxo-β-amyrin (almost 180°), which reversed the relative position of C-30 and C-29 **(Figure 5B)**. Furthermore, the influence of the residue on the regioselectivity in other P450s was also studied. CYP72A154 is another promiscuous P450 enzyme oxidizing 11-oxo-β-amyrin (**Figure S8**). The corresponding residue in CYP72A154 as L398 for CYP72A63 was predicted to be G397 and mutant CYP72A154 (G397I) was thus obtained. Using the substrate 11-oxo-β-amyrin, the mutant CYP72A154 (G397I) also showed 100% regioselectivity on C-29. The results provided evidence that residues on the equivalent locus of L398 on CYP72A63 were vital for the oxidation of C-29 on 11-oxo-β-amyrin for related P450s.

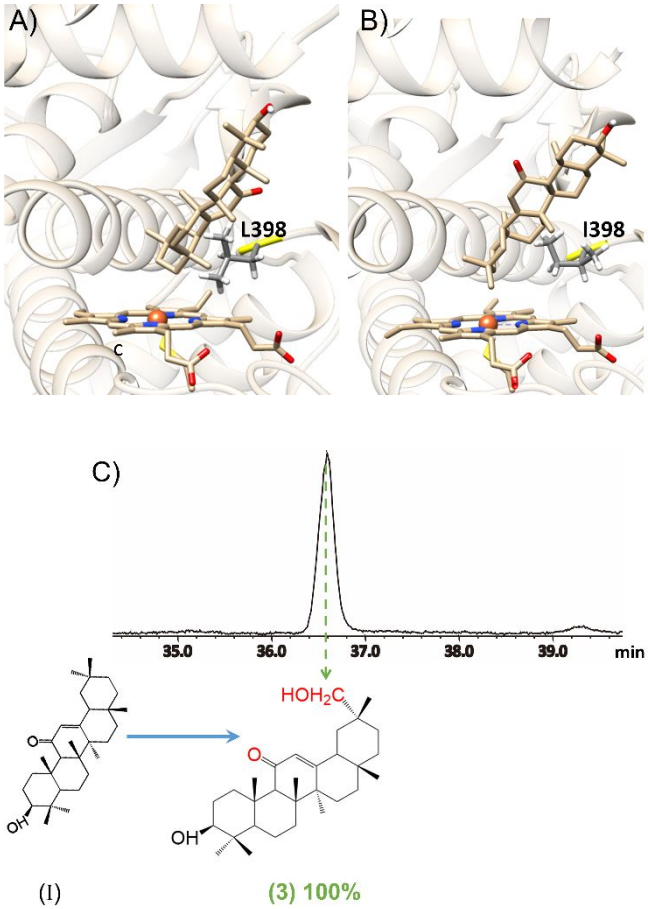


Figure 5. (A) Substrate pose in wild type CYP72A63, (B) Substrate pose in mutant

CYP72A63(L398I), (C) GC-MS spectrum showing the product distribution in yeast strain harboring CYP72A63 (L398I). (**I**) 11-oxo-β-amyrin, the substrate of CYP72A63 (3) 29-OH-11-oxo-β-amyrin, and the % means the percentage of the total oxidized products of 11-oxo-β-amyrin.

Enhancing the catalytic activity of CYP72A63 mutants by refining proton delivery

The mutant T338S discussed in the above section proves to be useful to improve the regioselectivity and consecutive oxidative ability of CYP72A63. Additionally, the total amount of different oxidation products of 11-oxo-amyrin increased to about 31.8 mg/L, 2.2 folds to 14.5mg/L for the wild type. The expression level was quantified by fusing the wild type *CYP72A63* and *CYP72A63 (T338S)* genes with *eGFP*, and there was no obvious difference in the expression level **(Figure S9)**.

We further quantified fermentation products of yeast with WT CYP72A63 or CYP72A63 (T338S) at different hours (24h, 48h, 72h and 120h) (**Figure S5**). The product distribution was similar at different times, and the mutant CYP72A63 (T338S) produced more oxidation product at any time, as well as that it had a higher ratio of glycyrrhetinic acid. This indicated CYP72A63(T338S) has enhanced oxidation activity.

To investigate the molecular basis, molecular oxygen was manually placed into the structure model with substrate 1.65 Å from the Fe, with a Fe–O–O angle of 135° and an O–O bond length of 1.26 Å.[41] The threonine hydroxyl points away from the molecular oxygen in the wild type CYP72A63 (**Figure 6A**). However, in the mutant T338S, the serine hydroxyl pointed to the active pocket and located beyond the heme bound molecular oxygen (**Figure 6B**). The minor variations led to refined protonation network by making the interaction between alcohol and heme bound O_2 much easier. Hence, more rapid formation of compound I and increased efficiency of catalysis was achieved. [42-43]

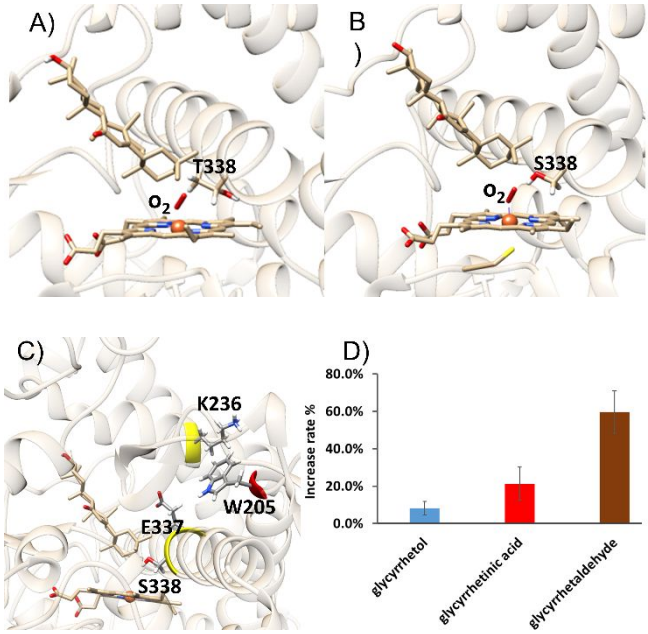


Figure 6. (A) The binding pattern of 11-oxo-β-amyrin in wild type CYP72A63 with heme bonded oxygen molecule. According to the mechanism of P450, the proton delivery would occur after inducing a conformational change in the proton delivery network in the P450 active site, followed by the direct donation of protons from the hydroxyl of alcohol-acid pair to heme bound O_2 leading to the cleavage of the O-O bond and thus formation of compound I. More rapid formation of compound I would allow less time for side product formation and result in increased efficiency of catalysis. (B) The binding pattern of 11-OXO-β-amyrin in CYP72A63 (T338S) with heme bonded oxygen molecule, (C) Proton transfer pathway. The proton transfer pathway consisted of K236, E337 and S338, and W205 has a significant hindrance hampering the conformational change of K236 (D) The increase rate of rare triterpenoids through strengthening proton delivery by mutant W205A

According to the model, K236 played a key role in donating the proton from solvent to the alcohol-acid pair, and this assumption can be verified by the mutant K236A which almost abolished the activity. Then it was clear that the proton shuttle pathway consisted of K236, E337 and S338 **(Figure 6C)**, and the mutant T338A, K236A, E337Q/A almost abolished the activity of CYP72A63 indicating that the typical proton delivery pathway was necessary for its activity, and the restriction of the proton transfer hindered the catalytic property. To find out the inherent restrictions for proton delivery of CYP72A63, residues around the proton shuttle pathway were analyzed and the results revealed that, the bulky sidechain of W205 hindering the conformational changing of K236 **(Figure 6C)** could be a major factor restricting the proton transfer. Hence, mutation W205L/A/R with different hindered side chain and property was carried out. Electrostatic repulsion formed between

K236 and R205 would further restrict the conformational change of K236, which hinders the proton delivery. As a result, CYP72A63(T338S/W205R) has much lower activity with the alcohol intermediate as the main product with full regioselectivity **(Figure S10)**. Although the mutant W205L has no significant influence on the catalytic activity, W205A mutation relieves the hindrance caused by tryptophan and removes the restriction on proton delivery. With no metabolic and process optimization conducted, CYP72A63(T338S/W205A) showed the best enhancement of the glycyrrhetinic acid production with a 21.3% increase to 36.4±3.0mg/L, which is the highest titer reported so far.

Considering this problem, the mutant CYP72A63 (T338S/W205A) with refined proton delivery combined with the GuCPR2 was employed to enhance the production of glycyrrhetaldehyde. The glycyrrhetaldehyde production in engineered strain GA175 harboring CYP72A63 (T338S/W205A) increased by 59.6% **(Figure 6D)** compared with that harboring CYP72A63 (T338S). The proton delivery pathway was also optimized in CYP72A63 (L509I) through iterative mutant W205A to promote the glycyrrhetol production. Compared with the strain with CYP72A63 (L509I), the glycyrrhetol production of strain GA159 harboring the mutant CYP72A63 (L509I/W205A) was increased by 8.2% **(Figure 6D)**. These results indicated that proton delivery was important to the catalytic property of P450s. Relieving the restriction that hindering the proton delivery is an effective way to increase specific P450 activity and thus enhance the synthesis of target products.

CONCLUSIONS

In summary, we remodeled the promiscuous CYP72A63 to controllable catalysts based on computation-guided mutations of rationally selected residues. Through homology modeling and docking, the structure model aided rational design of the P450 enzyme for licorice triterpenoid biosynthesis. The accuracy of the structure model need to be further tested experimentally, especially through detailed *in vitro* characterization of the enzyme. We believe this approach will be useful for engineering other eukaryotic P450s.

By manipulating the substrate accessibility and orientation of the target substrate through mutations on the proximal active pocket, the mutant CYP72A63 (T338S) achieved thorough oxidation ability on C-30 of 11-oxo-β-amyrin to produce glycyrrhetinic acid specifically. By

decreasing the accessibility of glycyrrhetol to CYP72A63 through mutations on the distal active pocket, the chemoselectivity was modified leading to controllable production of the intermediate glycyrrhetol. For the controllable production of glycyrrhetaldehyde, the mutant CYP72A63 (T338S) was coupled with proper CPRs, and GuCPR2 showed the best product spectrum with 82.3% glycyrrhetaldehyde. Moreover, the key factor for synthesizing the isomer was confirmed, and 29-OH-11-oxo-β-amyrin was specifically synthesized by the mutant CYP72A63 (L398I). More importantly, the catalytic activity of the mutants was further improved by refining the proton delivery with increased production of rare licorice triterpenoids in yeast. Our study provided a powerful model for controlling the catalytic property of P450s.

SUPPORTING INFORMATION

The Supporting Information is available free of charge on the ACS Publications website.

Chemicals, mutagenesis, time-course experiments, sequence alignment, homology modeling, biotransformation, method of purification, analytic method, GC-MS chromatograms, and NMR analysis.

ACKNOWLEDGEMENTS

This work was supported by the Natural Science Foundation of China (No. 21736002, and 21606018), the National Key Research and Development Program of China (2018YFA0901800) and the National Science Fund for Distinguished Young Scholars (No.21425624). We are grateful for the kind help of Prof. Yingjin Yuan and Bingzhi Li from Tianjin University for offering us yeast strain SynV.

REFERENCE

1. Zhang, R. K.; Huang, X.; Arnold, F. H., Selective CH bond functionalization with engineered heme proteins: new tools to generate complexity. *Curr Opin Chem Biol* 2018, *49*, 67-75.
2. Thimmappa, R.; Geisler, K.; Louveau, T.; O'Maille, P.; Osbourn, A., Triterpene biosynthesis in plants. *Annu Rev Plant Biol* 2014, *65*, 225-57.
3. Green, M. T., Role of the axial ligand in determining the spin state of resting cytochrome P450. *J Am Chem Soc* 1998, *120* (41), 10772-10773.
4. Green, M. T., Evidence for sulfur-based radicals in thiolate compound I intermediates. *J Am*

Chem Soc 1999, *121* (34), 7939-7940.

5. Meunier, B.; de Visser, S. P.; Shaik, S., Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. *Chem Rev* 2004, *104* (9), 3947-3980.
6. Loskot, S. A.; Romney, D. K.; Arnold, F. H.; Stoltz, B. M., Enantioselective Total Synthesis of Nigelladine A via Late-Stage C-H Oxidation Enabled by an Engineered P450 Enzyme. *J Am Chem Soc* 2017, *139* (30), 10196-10199.
7. Zhou, H.; Wang, B.; Wang, F.; Yu, X.; Ma, L.; Li, A.; Reetz, M. T., Chemo- and Regioselective Dihydroxylation of Benzene to Hydroquinone Enabled by Engineered Cytochrome P450 Monooxygenase. *Angew Chem Int Ed Engl* 2019, *58* (3), 764-768.
8. Le-Huu, P.; Heidt, T.; Claasen, B.; Laschat, S.; Urlacher, V. B., Chemo-, Regio-, and Stereoselective Oxidation of the Monocyclic Diterpenoid β-Cembrenediol by P450 BM3. *ACS Catal* 2015, *5* (3), 1772-1780.
9. Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I., Structure and chemistry of cytochrome P450. *Chem Rev* 2005, *105* (6), 2253-2278.
10. Rittle, J.; Green, M. T., Cytochrome P450 compound I: capture, characterization, and CH bond activation kinetics. *Science* 2010, *330* (6006), 933-937.
11. Urlacher, V. B.; Girhard, M., Cytochrome P450 Monooxygenases in Biotechnology and Synthetic Biology. *Trends Biotechnol*. *In press*.
12. Rabe, K. S.; Gandubert, V. J.; Spengler, M.; Erkelenz, M.; Niemeyer, C. M., Engineering and assaying of cytochrome P450 biocatalysts. *Anal Bioanal Chem* 2008, *392* (6), 1059-73.
13. Pompei, R.; Flore, O.; Marccialis, M. A.; Pani, A.; Loddo, B., Glycyrrhizic acid inhibits virus growth and inactivates virus particles. *Nature* 1979, *281* (5733), 689.
14. Liu, X.; Zhang, L.; Feng, X.; Lv, B.; Li, C., Biosynthesis of Glycyrrhetinic Acid-3-O-monoglucose Using Glycosyltransferase UGT73C11 from Barbarea vulgaris. *Ind Eng Chem Res* 2017, *56* (51), 14949-14958.
15. Kiso, Y.; Kato, O.; Hikino, H., Assay methods for antihepatotoxic activity using peroxide-induced cytotoxicity in primary cultured hepatocytes1. *Planta Med* 1985, *51* (01), 50-52.
16. Park, H.-Y.; Park, S.-H.; Yoon, H.-K.; Han, M. J.; Kim, D.-H., Anti-allergic activity of 18β-glycyrrhetinic acid-3-O-β-D-glucuronide. *Arch Pharmacal res* 2004, *27* (1), 57.

17. Okamoto, H.; Yoshida, D.; Saito, Y.; Mizusaki, S., Inhibition of 12-O-tetradecanoylphorbol-13-acetate-induced ornithine decarboxylase activity in mouse epidermis by sweetening agents and related compounds. *Cancer lett* 1983, *21* (1), 29-35.

18. Mukhopadhyay, M.; Panja, P., A novel process for extraction of natural sweetener from licorice (Glycyrrhiza glabra) roots. *Sep Purif Technol* 2008, *63* (3), 539-545.

19. Russel, F. G.; van Uum, S.; Tan, Y.; Smits, P., Solid-phase extraction of 18β-glycyrrhetinic acid from plasma and subsequent analysis by high-performance liquid chromatography. *J Chromatogr B* 1998, *710* (1-2), 223-226.

20. Tikhomirova, K.; Lekar, A.; Borisenko, S.; Vetrova, E.; Borisenko, N.; Minkin, V., Synthesis of glycyrrhetinic acid by hydrolysis of licorice root extract in subcritical water. *Russ J Phys Chem B* 2010, *4* (7), 1125-1129.

21. Wang, L. J.; Geng, C. A.; Ma, Y. B.; Huang, X. Y.; Luo, J.; Chen, H.; Zhang, X. M.; Chen, J. J., Synthesis, biological evaluation and structure-activity relationships of glycyrrhetinic acid derivatives as novel anti-hepatitis B virus agents. *Bioorg Med Chem Lett* 2012, *22* (10), 3473-9.

22. Rozen, S.; Shahak, I.; Bergmann, E., Synthesis and reactions of 18β-glycyrrhetaldehyde. *Tetrahedron* 1973, *29* (15), 2327-2331.

23. Fanani, M. Z.; Fukushima, E. O.; Sawai, S.; Tang, J.; Ishimori, M.; Sudo, H.; Ohyama, K.; Seki, H.; Saito, K.; Muranaka, T., Molecular Basis of C-30 Product Regioselectivity of Legume Oxidases Involved in High-Value Triterpenoid Biosynthesis. *Front Plant Sci* 2019, *10*, 1520.

24. Zhu, M.; Wang, C.; Sun, W.; Zhou, A.; Wang, Y.; Zhang, G.; Zhou, X.; Huo, Y.; Li, C., Boosting 11-oxo-beta-amyrin and glycyrrhetinic acid synthesis in Saccharomyces cerevisiae via pairing novel oxidation and reduction system from legume plants. *Metab Eng* 2018, *45*, 43-50.

25. Seki, H.; Sawai, S.; Ohyama, K.; Mizutani, M.; Ohnishi, T.; Sudo, H.; Fukushima, E. O.; Akashi, T.; Aoki, T.; Saito, K.; Muranaka, T., Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell* 2011, *23* (11), 4112-23.

26. Sun, W.; Qin, L.; Xue, H.; Yu, Y.; Ma, Y.; Wang, Y.; Li, C., Novel trends for producing plant triterpenoids in yeast. *Crit Rev Biotechnol* 2019, *39* (5), 618-632.

27. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 2004, *25* (13), 1605-1612.

28. 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 Crystallogr* 1993, *26* (2), 283-291.

29. Sippl, M. J., Calculation of conformational ensembles from potentials of mena force: An approach to the knowledge-based prediction of local structures in globular proteins. *J Mol Biol* 1990, *213* (4), 859-883.

30. Sippl, M. J., Recognition of errors in three-dimensional structures of proteins. *Proteins* 1993, *17* (4), 355-362.

31. Du, L.; Dong, S.; Zhang, X.; Jiang, C.; Chen, J.; Yao, L.; Wang, X.; Wan, X.; Liu, X.; Wang, X.; Huang, S.; Cui, Q.; Feng, Y.; Liu, S. J.; Li, S., Selective oxidation of aliphatic C-H bonds in alkylphenols by a chemomimetic biocatalytic system. *Proc Natl Acad Sci U S A* 2017, *114* (26), E5129-E5137.

32. Li, S.; Chaulagain, M. R.; Knauff, A. R.; Podust, L. M.; Montgomery, J.; Sherman, D. H., Selective oxidation of carbolide C–H bonds by an engineered macrolide P450 mono-oxygenase. *Proc Natl Acad Sci U S A* 2009, *106* (44), 18463-18468.

33. Kyte, J.; Doolittle, R. F., A simple method for displaying the hydropathic character of a protein. *J Mol Biology* 1982, *157* (1), 105-132.

34. Shimoyama, Y.; Hirabayashi, K.; Matsumoto, H.; Sato, T.; Shibata, S.; Inoue, H., Effects of glycyrrhetinic acid derivatives on hepatic and renal 11otein. n. dation of carbolide C–H bonds by an engin*J Pharm Pharmacol* 2003, *55* (6), 811-817.

35. Li, G.; Yao, P.; Gong, R.; Li, J.; Liu, P.; Lonsdale, R.; Wu, Q.; Lin, J.; Zhu, D.; Reetz, M. T., Simultaneous engineering of an enzyme's entrance tunnel and active site: the case of monoamine oxidase MAO-N. *Chem Sci* 2017, *8* (5), 4093-4099.

36. Narayan, A. R.; Jimenez-Oses, G.; Liu, P.; Negretti, S.; Zhao, W.; Gilbert, M. M.; Ramabhadran, R. O.; Yang, Y. F.; Furan, L. R.; Li, Z.; Podust, L. M.; Montgomery, J.; Houk, K. N.; Sherman, D. H., Enzymatic hydroxylation of an unactivated methylene C-H bond guided

by molecular dynamics simulations. *Nat Chem* 2015, *7* (8), 653-60.

37. Lim, R. K.; Lin, Q., Bioorthogonal chemistry: recent progress and future directions. *Chem Commun (Camb)* 2010, *46* (10), 1589-600.

38. De Montellano, P. R. O., *Cytochrome P450: structure, mechanism, and biochemistry*. Springer Science & Business Media: 2005, 33.

39. Sagadin, T.; Riehm, J. L.; Milhim, M.; Hutter, M. C.; Bernhardt, R., Binding modes of CYP106A2 redox partners determine differences in progesterone hydroxylation product patterns. *Commun Biol* 2018, *1*, 99.

40. Batabyal, D.; Richards, L. S.; Poulos, T. L., Effect of Redox Partner Binding on Cytochrome P450 Conformational Dynamics. *J Am Chem Soc* 2017, *139* (37), 13193-13199.

41. Lu, Q.; Song, J.; Wu, P.; Li, C.; Thiel, W., Mechanistic Insights into the Directing Effect of Thr303 in Ethanol Oxidation by Cytochrome P450 2E1. *ACS Catalysis* 2019, *9* (6), 4892-4901.

42. Schlichting, I.; Berendzen, J.; Chu, K.; Stock, A. M.; Maves, S. A.; Benson, D. E.; Sweet, R. M.; Ringe, D.; Petsko, G. A.; Sligar, S. G., The catalytic pathway of cytochrome P450cam at atomic resolution. *Science* 2000, *287* (5458), 1615-1622.

43. Harris, D. L.; Loew, G. H., A role for Thr 252 in cytochrome P450cam oxygen activation. *J Am Chem Soc* 1994, *116* (26), 11671-11674.