

Discovery and engineering of bacterial P450s for C-14 hydroxylation in *ent*-kaurane diterpenoids

Received: 13 May 2025

Accepted: 19 February 2026

Cite this article as: Lin, X., Xiao, Z., Xu, X. *et al.* Discovery and engineering of bacterial P450s for C-14 hydroxylation in *ent*-kaurane diterpenoids. *Nat Commun* (2026). <https://doi.org/10.1038/s41467-026-70157-0>

Xiaoxu Lin, Zhixi Xiao, Xingwang Xu, Xiaowei Zhang, Xingming Pan, Chenxi Zhu, Chenxi He, Li Feng, Fang-Ru Li, Hui-Min Xu, Zhe Wang, Ninghua Tan & Liao-Bin Dong

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Discovery and engineering of bacterial P450s for C-14 hydroxylation in *ent*-kaurane diterpenoids

Xiaoxu Lin,^{1,§} Zhixi Xiao,^{1,§} Xingwang Xu,¹ Xiaowei Zhang,¹ Xingming Pan,¹ Chenxi Zhu,² Chenxi He,³ Li Feng,¹ Fang-Ru Li,¹ Hui-Min Xu,⁴ Zhe Wang,¹ Ninghua Tan,^{1*} and Liao-Bin Dong^{1*}

¹State Key Laboratory of Natural Medicines, School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing 211198, China

²Mudi Meng Honors College, China Pharmaceutical University, Nanjing 211198, China

³School of Pharmacy, China Pharmaceutical University, Nanjing 211198, China

⁴The Public Laboratory Platform, China Pharmaceutical University, Nanjing 211198, China

[§]These authors contributed equally: Xiaoxu Lin, Zhixi Xiao.

*Corresponding authors: nhtan@cpu.edu.cn (N.T.); ldong@cpu.edu.cn (L.-B.D.)

Abstract

Ent-kaurane diterpenoids (*ent*-KTs) represent a structurally diverse class of natural products renowned for their antitumor and anti-inflammatory bioactivities. The C-14 hydroxyl modification is crucial for enhancing their potency; however, introducing this functional group remains a considerable challenge. Here, we present a computational heme-guided site-specific (CHS) strategy to identify three bacterial P450s (CYP260A1, CYP105N1, and CYP154C5) for C-14 hydroxylation. Further computationally guided enzyme engineering and redox partner screening identify the CYP260A1 L162V variant paired with CamA/CamB, achieving an 83-fold increase in production titer and a yield of 84.2 mg/L of (14*R*,16*R*)-*ent*-kauran-14,16-diol (**2**) in *Escherichia coli*. Substrate scope test reveals functional groups affecting reactivity. Structure-activity relationship studies demonstrate the synergistic effect between the C-14 hydroxyl and C15–C16 Michael acceptor, resulting in a potent derivative (**27**) with strong cytotoxicity ($IC_{50}^{HCT116} = 1.4 \mu M$). This study demonstrates a framework combining CHS-guided P450 discovery and computational enzyme engineering to advance *ent*-KT modifications.

Introduction

Ent-Kaurane diterpenoids (*ent*-KTs) are a prominent class of natural products with over 1,500 known compounds¹. Characterized by a distinctive tetracyclic skeleton, *ent*-KTs are highly valued for their pharmaceutical potential, particularly due to their notable antitumor and anti-inflammatory properties^{2,3}. The biological activity of *ent*-KTs is primarily attributed to the α,β -unsaturated carbonyl group within the 14-*O*- α -methylene cyclopentanone fragment, which engages in Michael addition reactions with cellular thiols⁴ (Fig. 1a and Supplementary Fig. 1). Structure-activity relationship (SAR) studies have demonstrated that further modifications at the C-14 hydroxyl group, achieved through the incorporation of either amino acids or saturated acid derivatives, can significantly enhance the biological activities of *ent*-KTs^{5,6}. For example, the oridonin A derivative HAO472, which is decorated with L-alanine at C-14, exhibited improved solubility and increased anticancer efficacy, ultimately advancing to a Phase I clinical trial for acute myelogenous leukemia⁷ (Fig. 1a). These findings highlight the C-14 hydroxyl group as a key structural feature for *ent*-KTs drug development.

Introducing a hydroxyl group at the C-14 position of *ent*-KTs remains challenging³. Total synthesis, which allows for the incorporation of a hydroxyl group at the C-14 position early in the process, is the most commonly favored approach^{8–10} (Fig. 1b). However, this method requires a complex design involving at least 12 steps, making it both challenge and costly. Directed C-H activation, extensively explored for functionalizing complex molecules,¹¹ has not yet succeeded in targeting the C-14 position of *ent*-KTs due to the sterically hindered environment and lack of suitable electronic biases. In contrast, biocatalysis, particularly through P450s, provides a highly regio- and stereoselective platform for C–H bond functionalization in terpenoids¹². In the context of *ent*-KTs, bacterial P450s—BM3 MERO1 M177A, PtmO6, and PtmO5—have been characterized as efficient biocatalysts for site-specific oxidation at the C-2, C-7, and C-11 positions, respectively^{13–15}. CYP706V6, a C-14 hydroxylase from *Isodon rubescens* involved in the biosynthesis of oridonin A, has been recently characterized but is limited in application due to

poor bacterial expression and a conversion efficiency of less than 5% in tobacco plants¹⁶ (Fig. 1b). Therefore, identifying C-14 P450s with improved expression and catalytic efficiency is crucial for advancing *ent*-KT functionalization.

Traditional P450 mining strategies, such as targeting homologs of known enzymes or screening predicted P450s through gene cluster analysis, cannot predict reactivity and reaction sites in advance^{17,18}. Structure-based and molecular docking methods have been developed to address these limitations^{19–23}. Successful examples include using structure modeling and analysis workflow to identify enzyme, applying structure-based approaches for deaminase, and combining these techniques to screen imine reductases from databases^{19,21,22}. Despite these advances, limitations in homology models and computational workflows have hindered accuracy. Recent developments in protein structure determination, particularly AlphaFold 3, provide new opportunities to overcome these challenges^{24,25}.

In this study, we develop a computational heme-guided site-specific (CHS) strategy to streamline the discovery of P450 enzymes capable of selective C-14 hydroxylation in *ent*-KTs. By integrating structural databases with P450 catalytic insights, we identify three bacterial C-14 hydroxylases (CYP260A1, CYP105N1, and CYP154C5). Enzyme engineering further enhances the activity of CYP260A1, leading to a significant increase in product yield. Through substrate scope expansion and SAR analysis, we elucidate key structural features influencing *ent*-KT bioactivity (Fig. 1c).

Results and discussion

Developing a CHS strategy for mining bacterial C-14 P450s

To identify functional P450s for *ent*-KTs, we developed a systematic screening strategy based on structural and functional considerations. From the InterPro database containing over 556,000 P450 sequences (as of November 2024)²⁶, we first selected 225 bacterial P450s with high-quality crystal structures and potential for *Escherichia coli* expression^{24,27}. We then narrowed down candidates

using terpenoid-specific criteria: genome neighborhood analysis identified 24 P450s co-localized with terpene synthases²⁸, while literature review revealed 20 bacterial P450s with confirmed terpenoid-oxidizing activity (Fig. 2a). This combined approach yielded 44 candidate P450s for detailed analysis (Supplementary Fig. 2, Supplementary Method 1-2 and Supplementary Data 1).

To further evaluate the potential binding modes of these P450s, we performed molecular docking analysis. Docking bulky terpenoid substrates into P450 active sites presents unique challenges, particularly in accurately defining substrate orientation relative to the heme group²⁹. To address this challenge, we employed AutoDock with the Lamarckian genetic algorithm (LGA)³⁰, which couples a global genetic algorithm with local search and encodes local improvements back into the population. This approach has been widely adopted in P450 studies for its enhanced sampling efficiency when dealing with conformationally flexible active sites^{31,32}. To address this issue, the docking cavity was defined based on either known substrate-protein complexes or structural homology analysis, using a 4 Å radius around the substrate binding region. We then conducted docking experiments with (16*R*)-*ent*-kauran-16-ol (**1**) as the initial substrate, calculating binding energy scores for each candidate. The docking results showed that most candidates exhibited binding energies below -5 kcal/mol, with some reaching -10 kcal/mol, suggesting potential binding capability (Supplementary Data 2 and Supplementary Method 3). Based on these computational predictions, we selected the top 10 candidates with the lowest binding energies for further analysis (Supplementary Fig. 3 and Supplementary Data 2).

The catalytic mechanism of P450s involves the activation of C-H bonds, where oxidation typically occurs at carbon atoms positioned closest to the heme iron center. The optimal distance between the heme iron and the target carbon is generally between 3.5 Å and 5 Å³³. Among the top 10 candidates, the iron of the heme in five P450s interacted with ring C or D while others primarily bound to ring A or B (Supplementary Fig. 4). Among the ring C or D-interacting enzymes, four had optimal distances to target C-14: CYP260A1 showed the closest distance at 3.5 Å, followed

by CYP105N1 and CYP105D7 at 3.6 Å, while CYP154C5 showed closer interaction with C-17 at 3.7 Å than C-14 at 4.8 Å (Fig. 3a). None of these P450s has been previously reported to act on *ent*-KTs. CYP260A1, initially identified in *Sorangium cellulosum* for its ability to hydroxylate nootkatone, has been shown to hydroxylate various steroid substrates^{34,35}. CYP105N1, from *Streptomyces coelicolor*, is involved in the biosynthesis of coelibactin siderophore and is bioinformatically linked to an isoborneol synthase^{36,37}. CYP105D7, from *Streptomyces avermitilis*, catalyzes the conversion of 1-deoxypentalenic acid to pentalenic acid³⁸. CYP154C5, found in *Nocardia farcinica*, catalyzes α -hydroxylation at the C-16 position of steroids³⁹ (Supplementary Fig. 5).

Establishing an *E. coli* system for P450s screening

Although traditional *in vitro* screening methods for P450s offer a clean background, they require labor-intensive protein purification and expensive NADPH regeneration systems^{40,41}. Moreover, these methods are often ineffective for hydrophobic substrates like **1**. *In vivo* systems preserve the natural membrane environment and NADPH/NADP⁺ cycle essential for P450 activity⁴². Additionally, compound **1**, which is difficult to obtain from natural or synthetic sources, can be directly produced in *E. coli* through its biosynthetic pathway¹⁵. Therefore, an *in vivo* screening platform to improve the efficiency of P450 enzyme screening was developed.

The PhoN-IPK system, a truncated two-step phosphorylation pathway for producing C5 precursors dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP) from dimethylallyl alcohol (DMAA) and isopentenol (ISO), has been previously established^{43–45} (Supplementary Fig. 6). We previously integrated this system with geranylgeranyl diphosphate synthase (GGDPS), *ent*-copalyl diphosphate synthase (*e*CDPS), and *ent*-kaurene synthase (BjKS) to achieve 113 ± 7 mg/L of *ent*-kauran-16-ene (**5**)⁴³. Based on this diterpene overproduction platform, we introduced three modifications to establish an *in vivo* P450 screening platform⁴⁶. First, BjKS was replaced with (16*R*)-*ent*-kauran-16-ol synthase (*e*KS) from *Streptomyces* sp. NRRL S-1813 under the control of a strong T7 promoter to generate **1** (Fig. 3b, Supplementary Figs. 7–12, Supplementary Method 4-

8 and Supplementary Data 3–5). Second, all biosynthetic genes (*phoN*, *ipk*, *idi*, *ggdps*, *ecdps*, and *ptmT3*) were consolidated into a single plasmid to construct strain DL10501¹⁵, simplifying the original three-plasmid system (Fig. 3c, Supplementary Table 1). Third, RhFRed from *Rhodococcus* sp. NCIMB 9784 was introduced as the redox partner for P450 functionality in *E. coli*.

Experimental validation of C-14 hydroxylases

The four P450 genes (CYP260A1, CYP105D7, CYP154C5, and CYP105N1) were fused with RhFRed redox partner gene and introduced into strain DL10501, generating strains DL10502–DL10505. Under optimized fermentation conditions, GC-MS analysis of the products revealed new peaks with m/z values of 288, confirming successful oxidation by all four P450s (Supplementary Method 7-11 and Supplementary Figs. 13 and 14). Scale-up fermentation was then performed to obtain sufficient products for NMR structural characterization (Supplementary Figs. 15–28, Supplementary Data 6, Supplementary Table 2). In the ¹H NMR spectrum of **2**, a characteristic oxidized methine proton signal was observed at δ_{H} 4.32 ppm, which corresponds to the ¹³C NMR signal at δ_{C} 78.9 ppm, as confirmed by the HSQC experiment. Detailed analysis of the ¹H–¹H COSY spectrum revealed spin–spin coupling patterns indicative of the interaction between the methine proton (δ_{H} 4.32) and H-9 (δ_{H} 2.05). In the HMBC spectrum, this oxidized methine proton showed correlations with C-15 (δ_{C} 56.0) and C-16 (δ_{C} 80.2). Additionally, the ROESY spectrum revealed a key correlation with H₃-20 (δ_{H} 0.99). These data, along with the X-ray crystal structure analysis of **2** (Fig. 3b, Supplementary Table 3, Supplementary Data 7), collectively support the identification of CYP260A1 as a C-14 hydroxylase responsible for introducing a β -hydroxy group. CYP154C5 was shown to produce a C-17 oxidized product of **3**, with δ_{H} 4.25 and δ_{C} 66.9 ppm that match literature values⁴⁷. Due to the low conversion rates, the product of CYP105N1 and the minor product of CYP154C5 could not be isolated. However, their identical retention times and fragmentation patterns in GC-MS analysis, compared to **2**, suggest that they are also C-14 oxidation products (Supplementary Fig. 13). Unexpectedly, CYP105D7 was identified as a C-3 hydroxylase, rather than the predicted C-14 hydroxylase, as evidenced by

signals at δ_{H} 4.27 and δ_{C} 77.5 ppm in the ^1H and ^{13}C NMR spectra of **4**, which are consistent with reported values⁴⁸.

To understand why CYP260A1 selectively hydroxylates C-14 while CYP105D7 targets C-3, we performed 100 ns molecular dynamics (MD) simulations with **1** (Supplementary Method 12). In CYP260A1, the substrate maintained a stable C-14 orientation toward the heme, with the ligand root mean square deviation (RMSD) remaining stable and the Fe–C-14 distance consistently within 3.5–5.0 Å throughout the simulation (Fig. 4f; Supplementary Figs. 29–31). Conversely, in CYP105D7, the C-14 orientation destabilized after ~60 ns with RMSD fluctuations of 1–7 Å. Instead, the C-3 position became increasingly stable, with ligand RMSD reduced to 2–3 Å and the Fe–C-3 distance averaging ~0.3 Å shorter than the Fe–C-14 distance in CYP260A1 (Supplementary Figs. 32–37). These simulations demonstrate that distinct active-site dynamics govern the complementary regioselectivity of these enzymes.

We further validated the CHS strategy's reliability by testing the remaining six P450 candidates (CYP150A6, CYP175A1, CYP105AS1, CYP154C4, CYP158A, and P450cin) (Supplementary Fig. 4). None produced peaks matching the retention time of the C-14 hydroxylated product from CYP260A1, confirming they lack C-14 hydroxylation activity (Supplementary Fig. 38). These results validate both our computational screening approach and its application to discovering site-specific P450 hydroxylases for *ent*-KTs.

Redox partners screening

CYP260A1 initially exhibited a low hydroxylation yield of 1.6 ± 0.3 mg/L with the redox partner of RhFRed, suggesting the requirement for systematic optimization to enhance its production efficiency. Due to the absence of native redox partners in the CYP260A1 gene cluster, several exogenous partners were evaluated, including CamA/CamB from *Pseudomonas putida*, Fdr/Fdx from *Spinach oleracea*, BM3Red from *Bacillus megaterium*, and CYP116B46 reductase from

*Thermus thermophilus*⁴⁹ (Supplementary Data 6 and Supplementary Method 11 and 13). GC-MS analysis revealed that CamA/CamB significantly improved the yield to 10.2 ± 0.6 mg/L, while BM3Red resulted in a reduced yield of only 0.6 ± 0.2 mg/L (Fig. 4a). Consistent with these experimental observations, protein–protein docking indicated that CamA/CamB engages CYP260A1 through a stronger and more complementary interface than RhFRed (Supplementary Fig. 39 and Supplementary Method 14). Having improved the electron transfer efficiency with CamA/CamB, we next turned to enhancing CYP260A1’s inherent activity through active-site mutagenesis guided by computational modeling.

MM/PBSA-guided active-site mutagenesis

To enhance CYP260A1’s production efficiency, we implemented a two-tier computational strategy that integrated MM/PBSA (Molecular Mechanics Poisson-Boltzmann Surface Area)-guided residue selection with FoldX- and docking-based virtual mutagenesis to optimize the active site (Supplementary Method 3 and 15). By combining dynamic enzyme–substrate–solvent interaction analysis with static mutational docking, this approach surpassed traditional static pocket analyses, enabling more accurate identification of residues influencing substrate binding while minimizing the need for extensive experimental screening.

In brief, MD simulations were conducted on CYP260A1 with substrate **1** bound in its active site to comprehensively sample the conformational ensemble of the enzyme–substrate complex. Binding free energy contributions of the 19 residues within 4 Å of the substrate were evaluated via per-residue MM/PBSA decomposition in Poisson-Boltzmann (PB) and Generalized Born (GB)^{50,51} (Figs. 4b and 4c and Supplementary Fig. 40). Because PB and GB can diverge on solvent-exposed, flexible loops, we treated PB as primary and used GB only as a sensitivity check⁵². In PB analysis, five residues (L159, L162, L228, G278, V279) exhibited unfavorable contributions ($\Delta G_{\text{res}} > 0$ kcal/mol), suggesting they may impose strain or suboptimal interactions, while seven residues (F64, V163, L224, F277, V332, F333, Y376) showed minimal contributions ($\Delta G_{\text{res}} \approx 0$ kcal/mol).

(Figs. 4b and 4c). In the GB analysis, nine residues (L69, A74, S276, F277, G278, V279, L280, V332, F333) exhibited results entirely inconsistent with those obtained from PB analysis (Fig. 4c, Supplementary Method 16 and Supplementary Fig. 40), indicating potential instability. We excluded four residues from engineering based on their structural characteristics. L224 and Y376 were eliminated due to their distance ($> 6 \text{ \AA}$) from the substrate interaction site, while S276 and L280 were excluded because they reside on a flexible, solvent-exposed loop with a root mean square fluctuation (RMSF) value of $1.6 \text{ \AA}$, significantly higher than the ensemble average of $0.01 \text{ \AA}$ (Supplementary Fig. 31). For such flexible loop regions, GB per-residue estimates are inherently less stable than PB calculations⁵². The exclusion of these loop residues was validated experimentally. The L280A variant exhibited a severely reduced titer ($0.8 \pm 0.2 \text{ mg/L}$), while substitutions at position S276 with other substrates produced bidirectional outcomes, including product profile shifts and reduced overall titers (Supplementary Table 4)^{34,35}. Consequently, the remaining twelve residues (F64, L69, A74, L159, L162, V163, L228, F277, G278, V279, V332, F333)—comprising both unfavorable and unstable contributors—were selected for virtual saturation mutagenesis. Each was systematically substituted *in silico* using FoldX, and binding affinities of the resulting variants were assessed with AutoDock Vina^{30,53}. This strategy enabled thermodynamically guided residue selection and high-throughput variant screening. From 228 variants, twelve mutations (F64V, L69V, A74L, L159M, L162V, V163G, L228A, F277W, G278A, V279L, V332R, F333A) were identified, with five (A74L, L162V, V163G, L228A, F333A) significantly improving docking-derived binding energy ($\Delta\Delta G_{\text{bind}} > 0.3 \text{ kcal/mol}$ vs. wild-type; Supplementary Fig. 41 and Supplementary Table 5).

Experimental validation and catalytic impact

To experimentally validate the computational predictions, we constructed and expressed the twelve suggested variants (Supplementary Fig. 41, Supplementary Method 17 and Supplementary Data 3–5) and evaluated their catalytic performance via GC-MS analysis (Fig. 4d, Supplementary Table 4, and Supplementary Method 18–21). The results showed good agreement with the computational

forecasts. The L162V variant, which MM/PBSA analysis identified as relieving a significant destabilizing effect of the wild-type residue, yielded 84.2 ± 10.3 mg/L of **2**—a striking 8.3-fold increase over wild-type. This direct alignment between predicted binding energy improvement and enhanced catalytic output underscores the predictive power of our approach. Similarly, A74L, another mutation with favorable predicted energetics, delivered a 3-fold yield increase (35.3 ± 5.2 mg/L), while L228A had a modest effect (10.8 ± 0.2 mg/L). In contrast, variants such as F64V, L69V, L159M, F277W, G278A, V279L, and V332R, all predicted to offer minimal binding energy improvements, indeed showed low yields (1–10 mg/L), further reinforcing the computational-experimental concordance. V163G was an exception, as it did not improve production titer. This outcome likely resulted from local structural disruptions in the vicinity of residue L162, which the computational modeling did not fully capture (Fig. 4e and Supplementary Figs. 44 and 45).

To understand the structural basis for improved activity, we performed 100 ns MD simulations of the top-performing L162V and A74L variants. Both variants exhibit lower-amplitude fluctuations than WT, with dampened mobility in the 150–180 residue region and fewer sharp peaks (Supplementary Fig. 29). The L162V mutation further positioned the substrate closer to the heme iron (Fe–C-14 distance: 3.5–4 Å in L162V vs. 3.5–5 Å in wide-type), and its RMSD over 100 ns trajectories was lowest and most stable (Fig. 4f, and Supplementary Figs. 30, 31, 42, and 43). Collectively, these results demonstrate that MM/PBSA-guided engineering successfully identified mutations that enhance catalysis by optimizing substrate positioning and stability within the active site, highlighting the advantage of dynamics-based approaches over static structural analysis for enzyme optimization.

Expansion of the substrate scope

To explore the substrate specificity of CYP260A1 L162V, twenty structurally diverse *ent*-KTs were synthesized, incorporating functional groups such as acetyl, hydroxyl, carboxyl, and keto moieties (Supplementary Figs. 46–93, Supplementary Data 6 and Supplementary Tables 2–3). GC-MS analysis identified six active substrates (Fig. 6 and Table 1). While *in vivo* transformation of **1**

yielded 84.2 ± 10.3 mg/L of product **2**, *in vitro* conversion was limited to 44%, likely due to the poor solubility of low-oxidized diterpenes, a phenomenon consistent with previous studies^{41,54}. Similarly, compound **7** yielded 41.7 mg/L *in vivo* but only 35% conversion *in vitro* (Fig. 6, Table 1 and Supplementary Figs. 94–95). Using (15*S*)-*ent*-kauren-15-ol (**8**) as a substrate resulted in 9% conversion to product **26**, whereas replacing the C-15 hydroxyl with a keto group (**9**) abolished reactivity. (Fig. 6, Table 1, Supplementary Fig. 95 and Supplementary Table 2). Modifications on the A-ring demonstrated distinct reactivity patterns. Compounds **6**, **12**, and **24**, each bearing hydroxyl or keto functionalities on the A-ring, exhibited conversion rates of 30%, 30%, and 39%, respectively, with the resulting products tentatively assigned as putative C-14-oxidized derivatives (**28–30**) on the basis of GC–MS and docking results (Supplementary Figs. 94–96). However, introducing bulkier substituents, such as acetyl or carboxylic acid groups, resulted in a loss of activity (Supplementary Fig. 46). A preliminary cascade reaction was attempted by combining CYP105D7-catalyzed C-3 hydroxylation with CYP260A1 L162V-mediated transformation. Several multi-oxidized products (*m/z* 306 and 322) were identified via mass spectrometry (Supplementary Fig. 95). Despite the low efficiency, this dual-enzyme system highlights the potential of coordinating P450s for complex *ent*-KTs formation and lays the groundwork for future combinatorial biosynthesis efforts.

Structure–activity relationship of *ent*-KTs

Compound **27**, featuring a 14-*O*- α -methylene cyclopentanone moiety, was synthesized through enzymatic C-14 oxidation (Fig. 5a). Due to the poor aqueous solubility of *ent*-kaurane (**5**) *in vitro*, C-14 hydroxylated **7** was produced *in vivo* using *E. coli*, yielding a titer of 41.7 mg/L. Subsequent SeO₂ oxidation at C-15 followed by IBX oxidation introduced a ketone group, giving **27** in 84% yield over two steps. The C-14 hydroxyl group remained unaltered during oxidation, likely due to steric hindrance, facilitating a protection-free, three-step synthesis of 14-oxygenated bicyclo[3.2.1]octane-containing *ent*-KT **27**.

To investigate the SAR of *ent*-KTs, synthesized compounds and intermediates were evaluated for their cytotoxicity against HCT116 and LLC cancer cell lines (Table 2). Compound **1** showed negligible activity (*IC*₅₀ > 100 μ M), while modifications at specific positions enhanced

cytotoxicity to varying extents. C-14 hydroxylation (**2**, $IC_{50}^{LLC} = 30.4 \pm 1.3 \mu M$ and C-17 hydroxylation (**4**, $IC_{50}^{LLC} = 29.0 \pm 7.2 \mu M$) both yielded moderate activities, whereas C-3 hydroxylation (**3**, $IC_{50}^{LLC} = 78.6 \pm 7.2 \mu M$) had a weaker effect (Table 2). Introducing a C-15=C-16 Michael acceptor group (**9**, $IC_{50}^{HCT116} = 5.5 \pm 1.6 \mu M$) significantly improved activity, which was further improved by adding a C-14 hydroxyl in **27** ($IC_{50}^{HCT116} = 1.4 \pm 0.7 \mu M$) (Table 2). Reducing the Michael acceptor in **27** to form **26** ($IC_{50}^{HCT116} = 49.4 \pm 4.8 \mu M$) led to a notable decline in activity, again highlighting the significance of the Michael acceptor for bioactivity. Compound **27** exhibited 25-fold greater potency than *cis*-platin and 8-fold higher than oridonin A ($IC_{50} = 11.6 \mu M$) against HCT116 cells⁵⁵ (Table 2). Dose–response analysis demonstrated that compound **27** nearly completely inhibited the target at 20 μM , a concentration below its aqueous solubility limit (33 μM) (Fig. 5b, Supplementary Method 22, Supplementary Fig. 97 and Supplementary Table 6). Mechanistic experiments confirmed that **27** induces apoptosis in HCT116 cells. Western blot analysis showed dose-dependent increases in cleaved Caspase-3 and cleaved PARP following treatment with **27** (Fig. 5c). Consistently, Annexin V/PI flow cytometry revealed a significant rise of late apoptosis (from 6.83% (control) to 39.1%) and early apoptosis (from 3.16% (control) to 30.2%), with a clear dose-response at 10 μM (Fig. 5d). Together, these results establish caspase-dependent apoptosis as the primary mode of action of **27** and suggest the synergistic contribution of the C-14 hydroxyl group and the Michael acceptor to the bioactivity of *ent*-KTs.

In summary, this study highlights the successful integration of computational and experimental approaches for the targeted discovery and optimization of C-14 hydroxylases in *ent*-KTs. By employing a CHS strategy, we identified three bacterial P450s—CYP260A1, CYP105N1, and CYP154C5—capable of catalyzing the challenging C-14 hydroxylation. Through enzyme engineering and redox partners screening, the CYP260A1 L162V variant exhibited an 83-fold improvement in production titer, significantly enhancing the production of **2**. The substrate scope expansion revealed specific functional groups that influence reactivity and selectivity, offering deeper insights into P450-catalyzed modifications. Furthermore, SAR studies underscored the importance of C-14 hydroxylation in conjunction with the C15–C16 Michael acceptor group for enhanced bioactivity. Compound **27** displayed potent cytotoxicity, highlighting its therapeutic potential. This research not only advances biocatalytic strategies for site-specific *ent*-KT functionalization but also provides a blueprint for future enzyme engineering and combinatorial

biosynthesis efforts. The integration of computational screening, enzyme optimization, and chemoenzymatic synthesis offers a robust platform for the efficient production of structurally diverse and bioactive terpenoids.

Methods

Plasmid construction and transformation

Genes were synthesized by General Biosystems (Anhui, China) or amplified from genomic DNA of *Kitasatosporia griseola* DSM43859, *S. sp.* NRRL S-1813, and *E. coli*. Vectors (pET-28a(+) and pACYCDuet-1) were linearized by restriction digestion (*Nde*I, *Bam*HI, *Xho*I, *Hind*III) and assembled with gene fragments via homologous recombination using ClonExpress II (Vazyme). Recombinant plasmids were transformed into chemically competent *E. coli* DH5 α or BL21(DE3) using a heat-shock method (42 °C, 45 s). Transformants were selected on LB agar with appropriate antibiotics and confirmed by colony PCR and DNA sequencing (Supplementary Data 3 and 5).

Protein expression and purification

Expression plasmids were transformed into *E. coli* BL21(DE3), and cultures were grown at 37 °C in LB medium containing kanamycin or chloramphenicol (25–50 μ g/mL) supplemented with trace metal mix (a solution providing essential metal ions required for cell growth of 50 mM FeCl₃, 20 mM CaCl₂, 10 mM MnSO₄, 10 mM ZnSO₄, 2 mM CoSO₄, 2 mM CuCl₂, 2 mM NiCl₂, 2 mM Na₂MoO₄, 2 mM H₃BO₃). Protein expression was induced at OD₆₀₀ 0.6 by adding IPTG (0.1–0.25 mM), DMAA (6–10 mM), and 5-aminolevulinic acid (0.5 mM), followed by incubation at 18 °C for 20 hours. Cells were harvested, resuspended in lysis buffer (50 mM Tris, pH 8.0, 300 mM NaCl, 10% glycerol), and lysed by sonication (60% amplitude, 1 s on/4 s off for 4 min). His-tagged proteins were purified using HisTrap affinity chromatography (GE Healthcare), eluted with an imidazole gradient, desalted via Hitrap columns, and concentrated using Amicon Ultra-15 concentrators. Protein concentrations were determined spectrophotometrically (Nanodrop 300), and purity confirmed by SDS-PAGE (Supplementary Fig. 98).

Fermentation and metabolite extraction

Engineered strains (DL10501–DL10527) were cultured in LB medium supplemented with antibiotics and trace metals. Induction was performed as described above, with cultures fermented

at temperatures of 18–25 °C and shaken at 200 rpm for 3–5 days. Cells were harvested, and metabolites extracted with acetone. Extracts were filtered (0.22 µm), concentrated, and purified by silica gel chromatography prior to GC-MS analysis. Data are presented as mean ± SD ($n = 3$, biological replicates).

GC-MS analysis

Extracted metabolites were analyzed using an EXPEC 5231 GC-MS system (Spectra Tech) equipped with an Agilent HP-5 column (30 m × 0.25 mm, 0.25 µm film thickness). The initial oven temperature was set at 150 °C (held 2 min), ramped to 300 °C at 10 °C/min, and held for an additional 3 min. Injection (1 µL, splitless mode) was performed with helium carrier gas (1.5 mL/min). Electron ionization (EI) at 70 eV was used, and data were analyzed with MassHunter software.

Enzymatic assays

CYP450-RhFRed enzymatic reactions were performed in 1 mL reaction volumes containing 40 µM CYP450-RhFRed, 10 µM Opt13, 100 µM substrate, 1 mM NADP⁺, and 100 mM Na₂HPO₃·H₂O in Tris-HCl buffer (pH 8.0). Incubations were carried out at 30 °C for 20 hours, followed by extraction with ethyl acetate. Product analysis was conducted by GC-MS or LC-MS. For CYP260A1 (wild type, L162V, and A74L), efforts to determine steady-state kinetic parameters by GC-MS/MS were impeded by the extremely low solubility and poor MS ionization efficiency of substrate **1**. Accordingly, catalytic performance was instead quantified on the basis of product titer and endpoint conversion (Supplementary Table 7; Supplementary Figs. 44, 45, and 99). The relative conversion rate was calculated using the following equation:

$$\text{Relative conversion rate} = (\text{Product peak area} / (\text{Substrate} + \text{Product peak area})) \times 100\% \quad (1)$$

Bacterial P450 library construction for terpene oxidation

A comprehensive virtual library of bacterial P450s capable of oxidizing terpenes was constructed systematically. Initially, all P450 sequences were retrieved from the InterPro and UniProt databases. These sequences were filtered to include only bacterial P450s with known crystal structures. The dataset was further refined using taxonomy filters to ensure bacterial origin.

Terpene oxidation capability was determined by reviewing literature, analyzing UniProt annotations, and examining biosynthetic gene clusters using the Genome Neighborhood Tool (<https://efi.igb.illinois.edu/efi-gnt>)²⁸.

Molecular docking studies

Docking analyses were performed using AutoDock 4.2 software³⁰. Protein crystal structures were obtained from the Protein Data Bank (PDB, <https://www.rcsb.org/>), prepared by removing water molecules, adding Kollman charges, and polar hydrogens using AutoDockTools (ADT)²⁹. Ligand structures were optimized with Gasteiger charges. Grid boxes encompassed the entire enzyme active sites (0.375 Å spacing). Docking employed the Lamarckian genetic algorithm (population size: 150, maximum evaluations: 2,500,000, generations: 27,000) with 100 docking runs per ligand. Pose generation used AutoDock's Lamarckian Genetic Algorithm (LGA), which couples a global genetic search with local minimization; the "Lamarckian" inheritance of local improvements enhances pose diversity and convergence, particularly for bulky, flexible ligands at fixed evaluation budgets⁵⁶.

Molecular dynamics simulations

MD simulations were conducted with GROMACS 2022.2 and AMBER99SB force field⁴⁷. Initial enzyme-substrate complexes were solvated in TIP3P water with neutralization and ionic strength adjustment (0.15 M NaCl). Systems underwent energy minimization, followed by equilibration under NVT and NPT conditions. Production simulations ran for 100 ns with a 2 fs timestep, using LINCS constraints and Particle Mesh Ewald (PME) for electrostatics. Data are expressed as mean $\pm$ standard deviation (SD) from three replicates. Binding energies were computed using gmx_MMPBSA, employing Generalized Born (GB) and Poisson-Boltzmann (PB) methods to determine contributions from key residues^{50,51}. MM/PBSA calculations were performed on MD ensembles using MMPBSA.py to obtain per-residue energy decompositions. This method was chosen because it provides residue-level electrostatic and van der Waals contributions at modest computational cost, suitable for screening large variant libraries⁵⁷. PB was used as the primary model and GB as a sensitivity check; residues showing strong PB–GB discrepancies, typically on flexible solvent-exposed loops, were excluded from mutagenesis. Binding energies were interpreted qualitatively for residue prioritization, not as predictors of catalytic turnover.

Cell lines and culture

Human colorectal cancer cells, HCT116 (RRID: CVCL_0291), were obtained from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). All media contained 10% FBS (Gibco) and 1% (v/v) penicillin–streptomycin (Invitrogen). The cells were cultured at 37 °C in a humidified incubator with a 5% (v/v) CO₂ atmosphere.

Western blotting

Cells treated with **27** were lysed on ice in lysis buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% Triton X-100, 1 mM EDTA) supplemented with protease and phosphatase inhibitor cocktails (Roche) for 15 min. Lysates were clarified by centrifugation (12,000 × g, 10 min, 4 °C) and quantified using the Detergent-Compatible Bradford assay (Beyotime). Equal protein amounts (20–30 µg) were mixed with 4× SDS loading buffer, boiled at 95 °C for 5 min, resolved by 10–12% SDS-PAGE, and transferred to PVDF membranes (Millipore; wet transfer, 100 V, 90 min). Membranes were blocked in 5% nonfat milk in TBST (TBS with 0.1% Tween-20) for 2 h at room temperature, incubated with primary antibodies (1:1000–1:2000; GAPDH, PARP, cleaved PARP, Caspase-3) overnight at 4 °C, washed three times in TBST (10 min each), and then incubated with HRP-conjugated secondary antibodies (1:5000) for 2 h at room temperature. Signals were developed with ECL (Tanon). Band intensities were background-subtracted, normalized to GAPDH, and reported as mean ± SEM ($n \geq 3$ independent biological experiments).

Flow cytometry

Apoptosis was analyzed by an Annexin V/PI staining kit (Yeesen). The **27**-treated cells were collected and washed with cold PBS. Then, the cells were incubated with Annexin V-FITC and PI for 15 min at 4 °C in the dark. The stained cells were subjected to the Invitrogen Attune NxT flow cytometer (Thermo Fisher Scientific), and the data were analyzed by FlowJo V10.

Data availability

All the data supporting the findings of this study are available in the main text, the Supplementary Information, or from the authors upon request. The X-ray crystal structures of **2** and **8** are available

at Cambridge Crystallographic Data Centre under identifier XUPSUA [<https://www.ccdc.cam.ac.uk/structures/Search?Ccdcid=2432865&DatabaseToSearch=Published>] and XUPTAH [<https://www.ccdc.cam.ac.uk/structures/Search?Ccdcid=2432866&DatabaseToSearch=Published>], respectively. Source data are provided with this paper.

References

1. Liu, M., Wang, W.-G., Sun, H.-D. & Pu, J.-X. Diterpenoids from *Isodon* species: an update. *Nat. Prod. Rep.* **34**, 1090–1140 (2017).
2. Sun, H.-D., Huang, S.-X. & Han, Q.-B. Diterpenoids from *Isodon* species and their biological activities. *Nat. Prod. Rep.* **23**, 673–698 (2006).
3. Zhao, X., Cachera, B., Hu, Q. & Ma, D. Recent advances in the synthesis of *ent*-kaurane diterpenoids. *Nat. Prod. Rep.* **39**, 119–138 (2022).
4. Kiely-Collins, H., Winter, G. E. & Bernardes, G. J. L. The role of reversible and irreversible covalent chemistry in targeted protein degradation. *Cell Chem. Biol.* **28**, 952–968 (2021).
5. Ding, Y. *et al.* Discovery and development of natural product oridonin-inspired anticancer agents. *Eur. J. Med. Chem.* **122**, 102–117 (2016).
6. Liu, J. *et al.* Identification of new potent anticancer derivatives through simplifying the core structure and modification on their 14-hydroxyl group from oridonin. *Eur. J. Med. Chem.* **231**, 114155 (2022).
7. Hu, X. *et al.* Effective enmein-type mimics of clinical candidate HAO472: design, synthesis and biological evaluation. *Eur. J. Med. Chem.* **171**, 169–179 (2019).
8. He, C., Hu, J., Wu, Y. & Ding, H. Total syntheses of highly oxidized *ent*-kaurenoids pharicin A, pharicin B, 7-*O*-acetylpsurata C, and psurata C: A [5+2] cascade approach. *J. Am. Chem. Soc.* **139**, 6098–6101 (2017).
9. Guo, J., Li, B., Ma, W., Pitchakuntla, M. & Jia, Y. Total synthesis of (-)-glaucocalyxin A. *Angew. Chem. Int. Ed.* **59**, 15195–15198 (2020).

10. Kong, L., Su, F., Yu, H., Jiang, Z., Lu, Y. & Luo, T. Total synthesis of (-)-oridonin: an interrupted nazarov approach. *J. Am. Chem. Soc.* **141**, 20048–20052 (2019).
11. Reetz, M. T., Qu, G. & Sun, Z. Engineered enzymes for the synthesis of pharmaceuticals and other high-value products. *Nat. Synth.* **3**, 19–32 (2024).
12. Zhang, K., El Damaty, S. & Fasan, R. P450 fingerprinting method for rapid discovery of terpene hydroxylating P450 catalysts with diversified regioselectivity. *J. Am. Chem. Soc.* **133**, 3242–3245 (2011).
13. Zhang, X. *et al.* Divergent synthesis of complex diterpenes through a hybrid oxidative approach. *Science* **369**, 799–806 (2020).
14. Dong, L.-B. *et al.* Cryptic and stereospecific hydroxylation, oxidation, and reduction in platensimycin and platencin biosynthesis. *J. Am. Chem. Soc.* **141**, 4043–4050 (2019).
15. Rudolf, J. D., Dong, L.-B., Manoogian, K. & Shen, B. Biosynthetic origin of the ether ring in platensimycin. *J. Am. Chem. Soc.* **138**, 16711–16721 (2016).
16. Sun, Y. *et al.* A chromosome-level genome assembly reveals that tandem-duplicated CYP706V oxidase genes control oridonin biosynthesis in the shoot apex of *Isodon rubescens*. *Mol. Plant* **16**, 517–532 (2023).
17. Lin, G.-M. & Voigt, C. A. Design of a redox-proficient *Escherichia coli* for screening terpenoids and modifying cytochrome P450s. *Nat. Catal.* **6**, 1016–1029 (2023).
18. Vanacek, P. *et al.* Exploration of enzyme diversity by integrating bioinformatics with expression analysis and biochemical characterization. *ACS Catal.* **8**, 2402–2412 (2018).
19. Huang, J. *et al.* Discovery of deaminase functions by structure-based protein clustering. *Cell* **186**, 3182–3195.e3114 (2023).
20. Lobb, B. & Doxey, A. C. Novel function discovery through sequence and structural data mining. *Curr. Opin. Struct. Biol.* **38**, 53–61 (2016).
21. Xu, K. *et al.* Structure-guided discovery of highly efficient cytidine deaminases with sequence-context independence. *Nat. Biomed. Eng.* **9**, 93–108 (2025).

22. Yu, Y. *et al.* Structure-based design of small imine reductase panels for target substrates. *ACS Catal.* **13**, 12310–12321 (2023).
23. Vasina, M. *et al.* Advanced database mining of efficient haloalkane dehalogenases by sequence and structure bioinformatics and microfluidics. *Chem Catal.* **2**, 2704–2725 (2022).
24. Taylor, R. & Wood, P. A. A Million crystal structures: the whole is greater than the sum of its parts. *Chem. Rev.* **119**, 9427–9477 (2019).
25. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
26. Paysan-Lafosse, T. *et al.* InterPro in 2022. *Nucleic Acids Res.* **51**, D418–D427 (2023).
27. Li, S., Du, L. & Bernhardt, R. Redox partners: function modulators of bacterial P450 enzymes. *Trends Microbiol.* **28**, 445–454 (2020).
28. Gerlt, J. A. *et al.* Enzyme function initiative-enzyme similarity tool (EFI-EST): a web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta, Proteins Proteomics* **1854**, 1019–1037 (2015).
29. Fan, J., Fu, A. & Zhang, L. Progress in molecular docking. *Quant. Biol.* **7**, 83–89 (2019).
30. Forli, S. *et al.* Computational protein–ligand docking and virtual drug screening with the autodock suite. *Nat. Protoc.* **11**, 905–919 (2016).
31. Santos, S. F. G., Bommareddy, R. R., Black, G. W., Singh, W. & Huang, M. The substrate specificity in the O -demethylation of 4-alkylguaiacols by cytochrome P450 AgcA_{P450}. *Catal. Sci. Technol.* **13**, 2070–2079 (2023).
32. Liu, Z. *et al.* Evolutionary interplay between sister cytochrome P450 genes shapes plasticity in plant metabolism. *Nat Commun* **7**, 13026 (2016).
33. Shahrokh, K., Orendt, A., Yost, G. S. & Cheatham III, T. E. Quantum mechanically derived amber-compatible heme parameters for various states of the cytochrome P450 catalytic cycle. *J. Comput. Chem.* **33**, 119–133 (2012).

34. Schiffrin, A., Litzenburger, M., Ringle, M., Ly, T. T. B. & Bernhardt, R. New sesquiterpene oxidations with CYP260A1 and CYP264B1 from *Sorangium Cellulosum* so Ce56. *ChemBioChem* **16**, 2624–2632 (2015).
35. Khatri, Y. *et al.* Structure-based engineering of steroidogenic CYP260A1 for stereo- and regioselective hydroxylation of progesterone. *ACS Chem. Biol.* **13**, 1021–1028 (2018).
36. Zhao, B. *et al.* Structural analysis of cytochrome P450 105N1 involved in the biosynthesis of the zincophore, coelibactin. *Int. J. Mol. Sci.* **13**, 8500–8513 (2012).
37. Lim, Y.-R. *et al.* Crystal structure of cytochrome P450 CYP105N1 from *Streptomyces coelicolor*, an oxidase in the coelibactin siderophore biosynthetic pathway. *Arch. Biochem. Biophys.* **528**, 111–117 (2012).
38. Ma, B. *et al.* Hydroxylation of steroids by a microbial substrate-promiscuous P450 cytochrome (CYP105D7): key arginine residues for rational design. *Appl. Environ. Microbiol.* **85**, e01530-19 (2019).
39. Bracco, P., Janssen, D. B. & Schallmeyer, A. Selective steroid oxyfunctionalisation by CYP154C5, a bacterial cytochrome P450. *Microb. Cell Fact.* **12**, 95-105 (2013).
40. Zhao, H. & Van Der Donk, W. A. Regeneration of cofactors for use in biocatalysis. *Curr. Opin. Biotechnol.* **14**, 583–589 (2003).
41. Rudolf, J. D. *et al.* Cytochrome P450-catalyzed hydroxylation initiating ether formation in platensimycin biosynthesis. *J. Am. Chem. Soc.* **140**, 12349–12353 (2018).
42. Ren, L. *et al.* Functional characterization of CYP81C16 involved in the tanshinone biosynthetic pathway in *Salvia Miltiorrhiza*. *Chin. J. Nat. Med.* **21**, 938–949 (2023).
43. Li, F.-R. *et al.* Efficient production of clerodane and *ent*-kaurane diterpenes through truncated artificial pathways in *Escherichia coli*. *Beilstein J. Org. Chem.* **18**, 881–888 (2022).
44. Du, W. *et al.* Microbe engineering to provide drimane-type building blocks for chiral pool synthesis of meroterpenoids. *Angew. Chem. Int. Ed.* **64**, e202419463 (2025).
45. Lund, S., Hall, R. & Williams, G. J. An artificial pathway for isoprenoid biosynthesis decoupled from native hemiterpene metabolism. *ACS Synth. Biol.* **8**, 232–238 (2019).

46. Xiao, Z. *et al.* Cytochrome P450-mediated skeleton rearrangement of taxadiene in an engineered *Escherichia coli* system. *Org. Lett.* **26**, 1640–1644 (2024).
47. Pacheco, A. G., Machado De Oliveira, P., Piló-Veloso, D. & Flávio De Carvalho Alcântara, A. ^{13}C -NMR data of diterpenes isolated from *Aristolochia* species. *Molecules* **14**, 1245–1262 (2009).
48. Tori, M., Arbiyanti, H., Taira, Z. & Asakawa, Y. Terpenoids of the liverwort *Fruuanoides denszfolia* and *Trocholejeunea sandvicensis*. *Phytochemistry* **32**, 335–348 (1993).
49. Chen, C.-C. *et al.* Molecular basis for a toluene monooxygenase to govern substrate selectivity. *ACS Catal.* **12**, 2831–2839 (2022).
50. Valdés-Tresanco, M. S., Valdés-Tresanco, M. E., Valiente, P. A. & Moreno, E. gmx_MMPBSA: A new tool to perform end-state free energy calculations with GROMACS. *J. Chem. Theory Comput.* **17**, 6281–6291 (2021).
51. Miller, B. R. I. *et al.* MMPBSA.py: an efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **8**, 3314–3321 (2012).
52. Sun, H., Li, Y., Tian, S., Xu, L. & Hou, T. Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. *Phys. Chem. Chem. Phys.* **16**, 16719–16729 (2014).
53. Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F. & Serrano, L. The FoldX web server: an online force field. *Nucleic Acids Res.* **33**, W382–W388 (2005).
54. Wang, J. *et al.* Discovery of key cytochrome P450 monooxygenase (C20ox) enables the complete synthesis of tripterifordin and neotripterifordin. *ACS Catal.* **15**, 2690–2702 (2025).
55. Zhang, D. *et al.* ROS/JNK/c-Jun axis is involved in oridonin-induced caspase-dependent apoptosis in human colorectal cancer cells. *Biochem. Biophys. Res. Commun.* **513**, 594–601 (2019).
56. Morris, G. M. *et al.* Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639–1662 (1998).

57. Wang, E. *et al.* End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. *Chem. Rev.* **119**, 9478–9508 (2019).

Acknowledgments

The authors thank Prof. Shengying Li from Shandong University for kindly providing plasmids pET28a (+)-FdX and pET28a (+)-FdR and Prof. Xiao Zhang from Southwest University for *ent-kaura-16-en-19-oic* acid. This work is supported in part by the National Science Foundation of China Grant (82473823 to L.-B.D.), the Natural Science Foundation of Jiangsu Province for Distinguished Young Scholars (BK20240097 to L.-B.D.), the Jiangsu Provincial Major Science and Technology Special Project (BG2024046 to L.-B.D.), the Fundamental Research Funds for the Central Universities (2632025TD06 to L.-B.D.), and the Project Program of State Key Laboratory of Natural Medicines at China Pharmaceutical University (SKLNMZZ2024JS44 to L.-B.D.).

Author contributions

L.-B.D. and N.T. conceived the project and supervised the study; X.L., Z.X., and L.-B.D. designed the experiments; X.L., Z.X., X.X., C.Z., C.H., F.-R.L. and H.-M.X. performed the experiments; X.L. carried out the computational study; L.F. and Z.W. conducted the experiments related to biological activity; all authors analyzed and discussed the results; X.L., X.P., X.Z. and L.-B.D. wrote the manuscript with inputs from all co-authors.

Competing interests

The authors declare no competing interests.

Table 1. Substrate scope of CYP260A1. Various *ent*-KTs were tested as CYP260A1 substrates, with **2** and **7** achieving high *in vivo* yields, while others showed variable conversion rates.

*Indicates identification performed only by GC-MS.

Substrate	Product	Conversion (%)
1	2	44%, 84.2 mg/L <i>in vivo</i>
5	7	35%, 41.7 mg/L <i>in vivo</i>
6	28*	30%
8	26	9%
9	27	<1%
12	29*	30%
24	30*	39%

Table 2. Antitumor activity evaluation. Cytotoxicity assays against human colorectal tumor (HCT116) and Lewis lung carcinoma (LLC) cell lines identified **27** as the most potent derivative. *Cis-platin* was used as a reference drug for comparison. Data are expressed as mean $\pm$ standard deviation (SD) from three biological replicates. Compounds marked with two asterisks (**) indicate that precipitation occurred during bioactivity tests. Source data are provided as a Source Data file.

Cell line	IC ₅₀ (μ M)							
	1	2	3	4	9	26	27	<i>cis-platin</i>
HCT116	>100**	>100	>100	>100	5.5 $\pm$ 1.6	49.4 $\pm$ 4.8	1.4 $\pm$ 0.7	24.2 $\pm$ 3.8
LLC	>100	30.4 $\pm$ 1.3	78.6 $\pm$ 7.2	29.0 $\pm$ 7.2	11.0 $\pm$ 1.2	56.4 $\pm$ 8.1	2.1 $\pm$ 0.5	15.7 $\pm$ 4.1

Figure legends

Fig. 1. Strategies for C-14 hydroxylation of *ent*-kaurane diterpenoids (*ent*-KTs). **a** Representative bioactive *ent*-KTs. HAO472, which progressed to a Phase I clinical trial for acute myelogenous leukemia, highlights the significance of C-14 hydroxylation. **b** Existing C-14 hydroxylation strategies. Chemical synthesis requires 12–26 steps, while CYP706V6-mediated enzymatic hydroxylation suffers from poor *E. coli* expression and low efficiency. **c** Discovery and engineering of bacterial C-14 hydroxylases for *ent*-KT functionalization. This study integrates heme-guided site-specific (CHS) screening for bacterial C-14 hydroxylase discovery, computational-guided CYP260A1 engineering (84.2 mg/L yield), and SAR studies.

Fig. 2. Workflow of the C-14 hydroxylase mining and engineering. **a** CHS strategy utilization. A CHS approach screened 225 bacterial P450s, selecting 45 terpene-related enzymes for molecular docking. The top 10 candidates were ranked by binding energy, with four enzymes exhibiting optimal C-14 reaction distances, followed by *in vivo* characterization. **b** Enzyme engineering. Five redox partners were evaluated, and molecular dynamics (MD) simulations analyzed enzyme-substrate interactions. MM/PBSA analysis (4 Å) identified key residues, leading to 228 virtual saturation mutations for optimization. **c** SAR evaluation. The engineered enzyme facilitated a substrate scope screen (20 *ent*-KTs) and chemoenzymatic synthesis, yielding hydroxylated derivatives. A target product (**27**) exhibited potent cytotoxicity.

Fig. 3. Identification and screening of bacterial C-14 hydroxylases for *ent*-KTs functionalization. **a** Structural analysis of bacterial P450 candidates. Molecular docking reveals that CYP260A1, CYP105D7 and CYP105N1 position C-14 within 3.5–3.6 Å of the heme, while CYP154C5 interacts with both C-14 and C-17. **b** Two-module *in vivo* P450 screening system. Module I generates the *ent*-KT substrate (**1**) via a biosynthetic pathway using the C5 isoprenol substrates of dimethylallyl alcohol (DMAA) and isopentenol (ISO). Module II introduces candidate P450s (CYP260A1, CYP105N1, CYP154C5, and CYP105D7) for hydroxylation

screening, producing hydroxylated derivatives (2–4). **c** Optimized substrate production. Fermentation conditions were optimized by varying glycerol concentration, optical density (OD₆₀₀) at induction, and cultivation time, achieving a maximum yield of 520.8 ± 5.9 mg/L for **1**. Data are presented as mean $\pm$ SD ($n = 3$, biological replicates). **d** GC-MS analysis of hydroxylated products. CYP260A1 and CYP105N1 selectively hydroxylated C-14, while CYP154C5 catalyzed hydroxylation at both C-14 and C-17. In contrast, CYP105D7 produced a hydroxylated derivative at C-3. Source data are provided as a Source Data file.

Fig. 4. Optimization of CYP260A1 via redox partner screening and computational-guided enzyme engineering. **a** Redox partner screening. CYP260A1 activity was assessed with different redox partners, showing CamA/CamB significantly enhanced product yield compared to the original RhFRed system. Data are presented as mean $\pm$ SD ($n = 3$, biological replicates). **b** Active site analysis. MM/PBSA analysis identifies residues within 4 Å of CYP260A1's active site that negatively impact substrate binding. **c** Energy contribution analysis with Poisson-Boltzmann (PB). Residues with positive values negatively impacted catalysis, while those with negative values improved substrate binding. **d** Experimental validation of mutations. The L162V variant exhibited the highest production titer, increasing product yield to 84.2 ± 10.3 mg/L, an 8.3-fold improvement over the wild-type. Error bars indicate the standard deviation of three independent biological replicates. Statistical analysis was performed using a one-way ANOVA across all genotypes, followed by Tukey's two-sided multiple-comparisons test. **e** Active-site view highlighting two key variants (L162V and A74L) of CYP260A1, with blue and red denoting the side chains of the wild-type and mutated residues, respectively. **f** Distance between the C-14 atom of substrate **1** and the heme iron center over a 100 ns MD simulation. The black, blue, and red traces represent the distances for wild-type, A74L, and L162V, respectively. Source data are provided as a Source Data file.

Fig. 5. Enzymatic and chemoenzymatic synthesis of 14-oxygenated *ent*-KTs and bioactivity evaluation. **a** Chemoenzymatic synthesis of **27** featuring a typical 14-*O*- α -methylene cyclopentanone moiety. **b** Cell viability assay showing concentration-dependent inhibition of HCT116 cell growth by **27**. Data are presented as mean $\pm$ SD ($n = 3$, biological replicates). **c** Expression of apoptosis-related proteins in HCT116 cells treated with **27** at the indicated concentrations for 24 h, as determined by Western blotting ($n = 3$ biological replicates). **d** Annexin V/PI flow cytometry analysis of apoptosis in HCT116 cells treated with **27** at the indicated concentrations for 24 h. Quadrants Q2 (Annexin V⁺/PI⁺) indicate late apoptotic cells, and the percentage is shown in red. Compound **27** induced a dose-dependent increase in apoptotic cell populations compared with the vehicle control. Source data are provided as a Source Data file.

Fig. 6. Major substrates synthesized showed reactivity toward CYP260A1. R¹–R⁴ substitutions are indicated; CYP260A1-catalyzed hydroxylation *in vivo* or *in vitro* is shown (red OH).

Editor's Summary

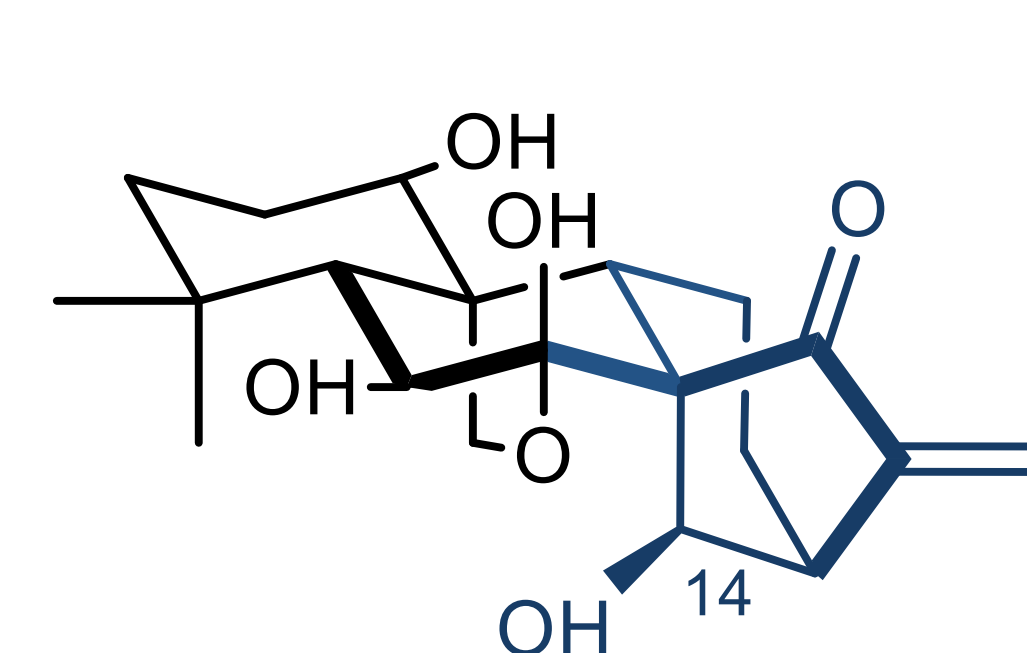
C-14 hydroxyl group is critical for the function of *ent*-kaurane diterpenoids (*ent*-KTs). Here, the authors report discovery and engineering of bacterial P450s for C-14 hydroxylation in *ent*-KTs.

Peer review information: *Nature Communications* thanks Carlos Acevedo-Rocha, Reuben Peters and Yong Wang for their contribution to the peer review of this work. A peer review file is available.

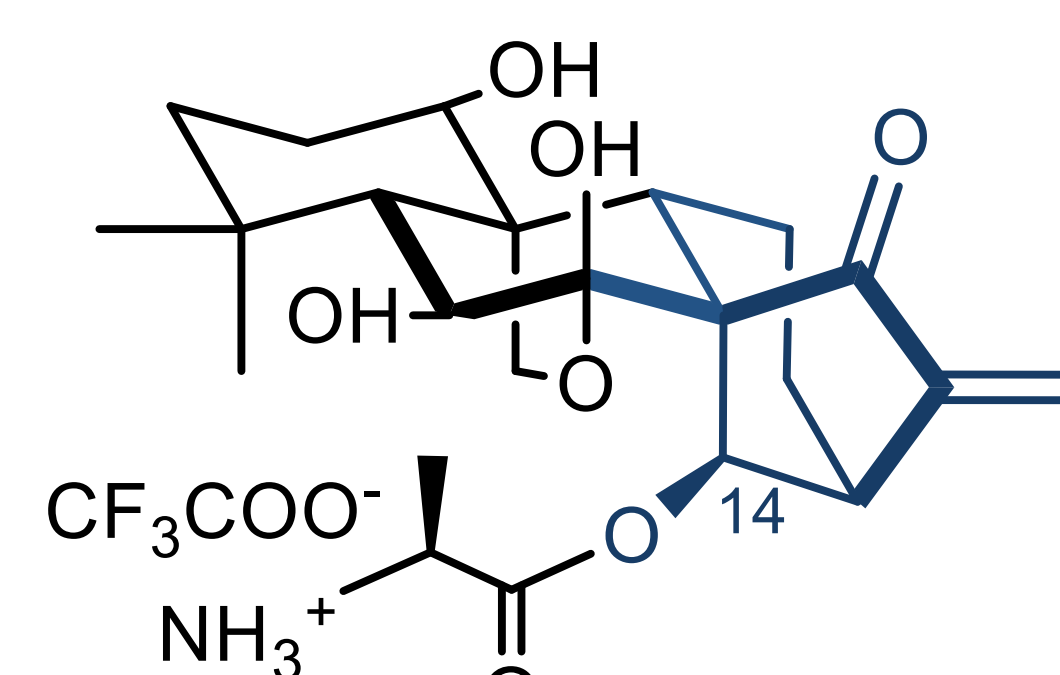
ARTICLE IN PRESS

Representative bioactive ent-KTs

a



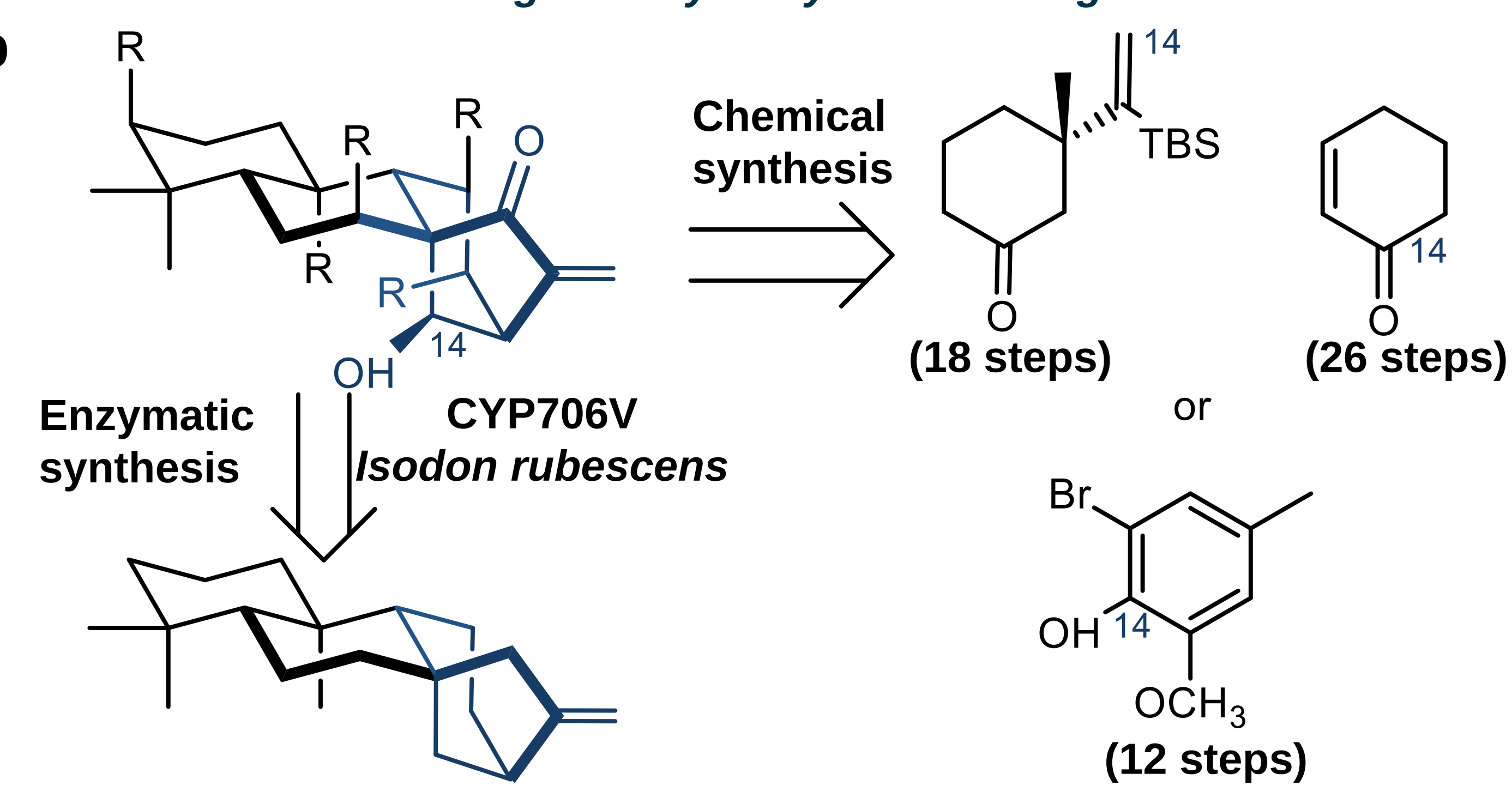
Oridonin A
 $IC_{50}^{HCT116} = 11.6 \mu M$



HAO472
 Phase I clinical trial for
 acute myelogenous leukemia

Existing C-14 hydroxylation strategies

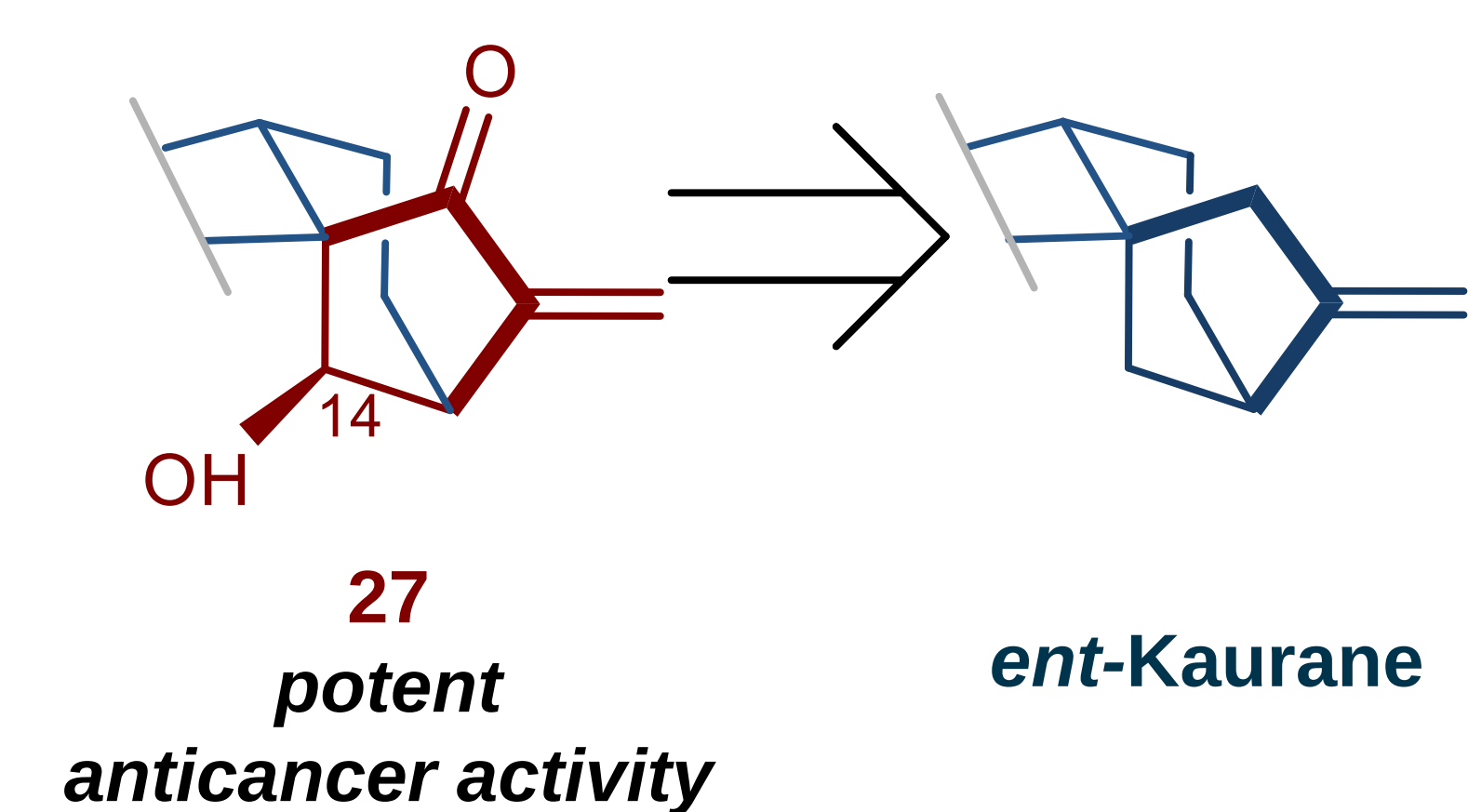
b



- **poor** expression in *E. coli*
- **poor** catalytic activity
- **lengthy** synthetic routes
- **low** yields

Discovery and engineering of bacterial C-14 hydroxylases

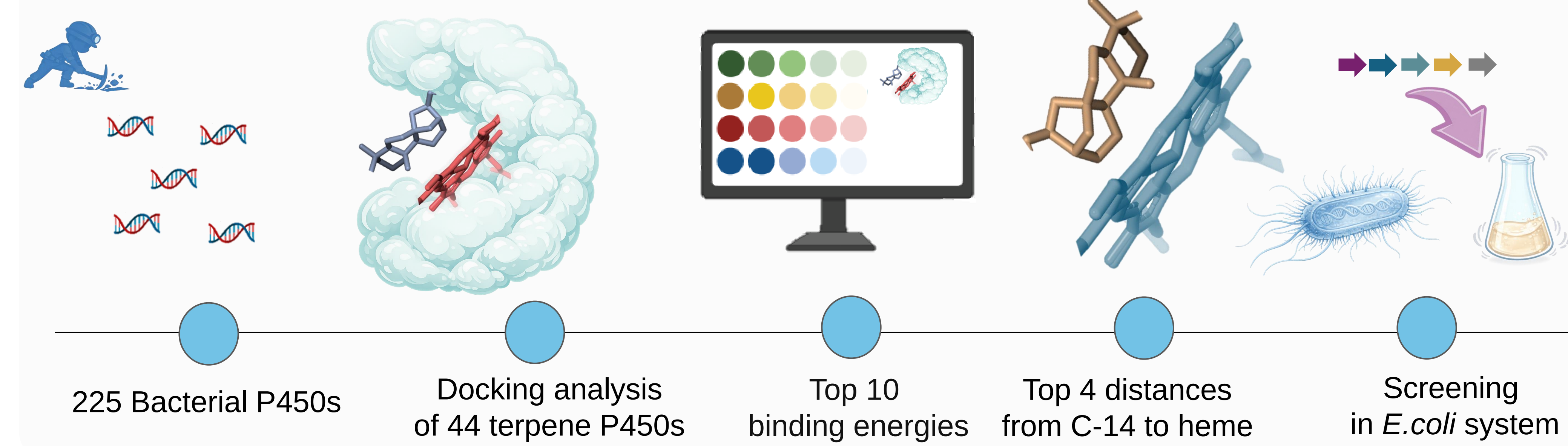
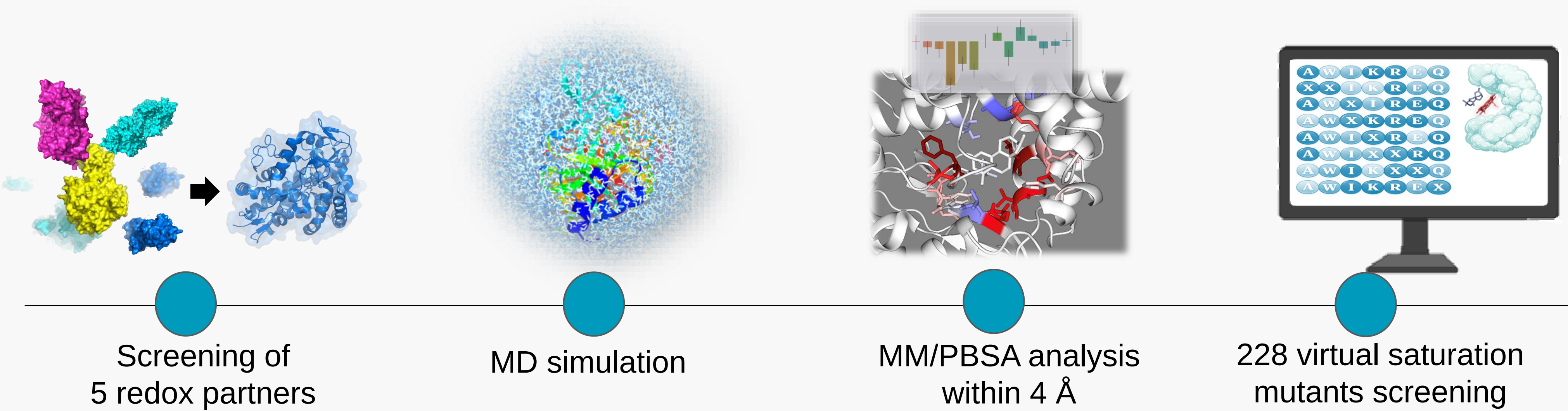
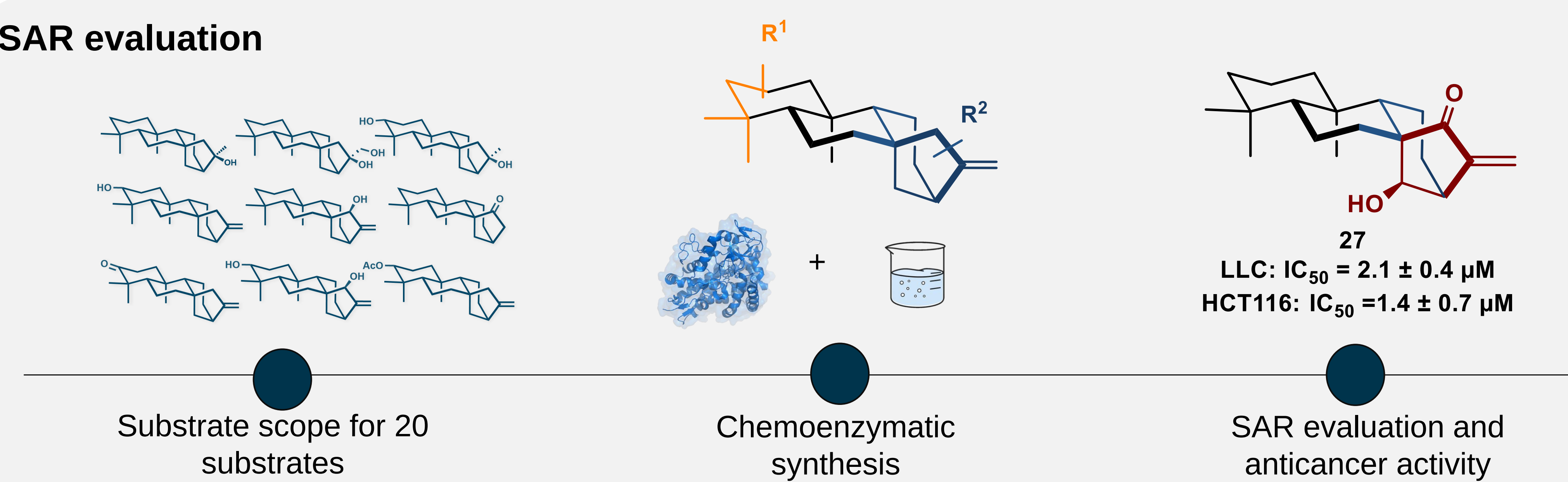
c

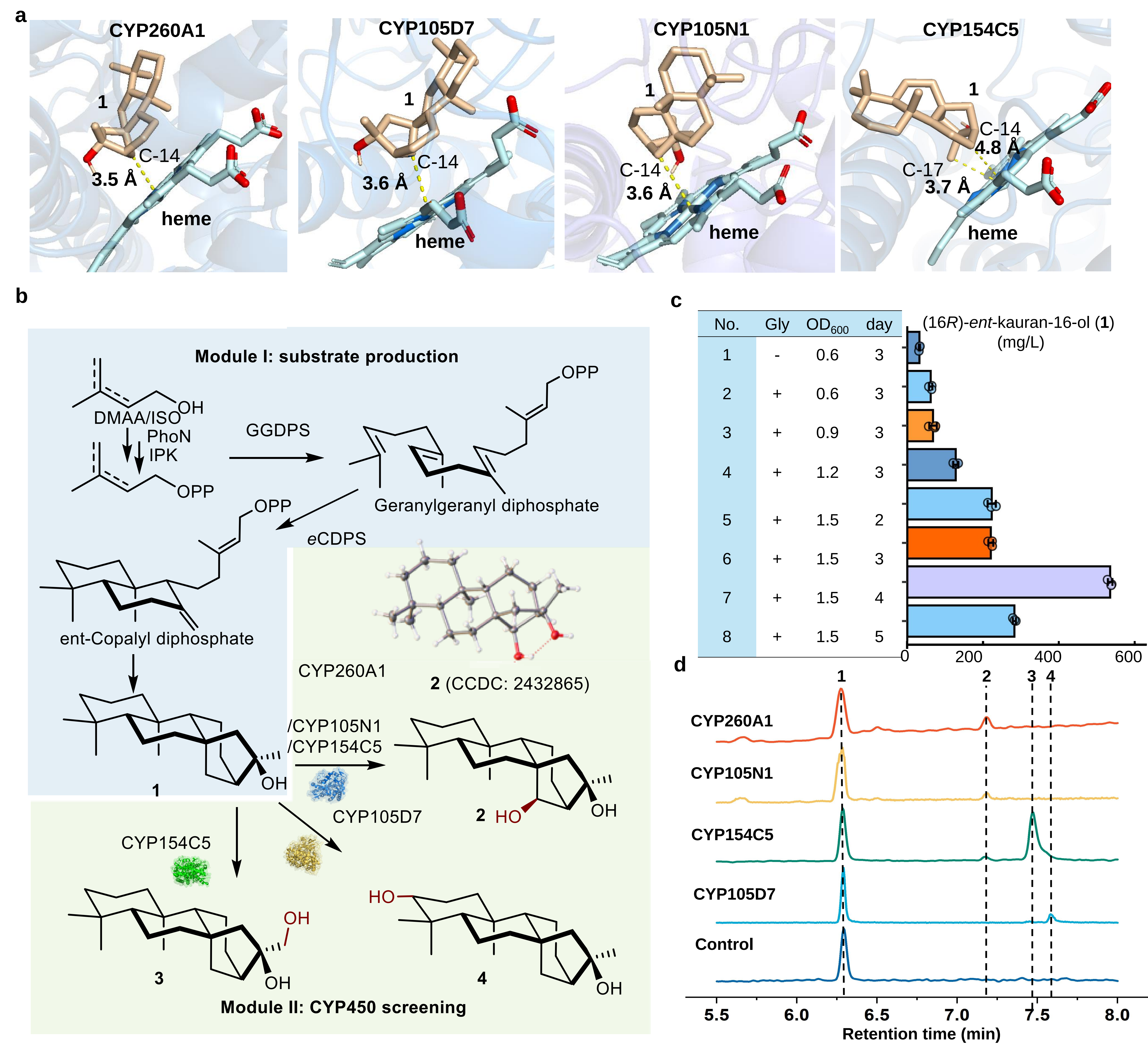


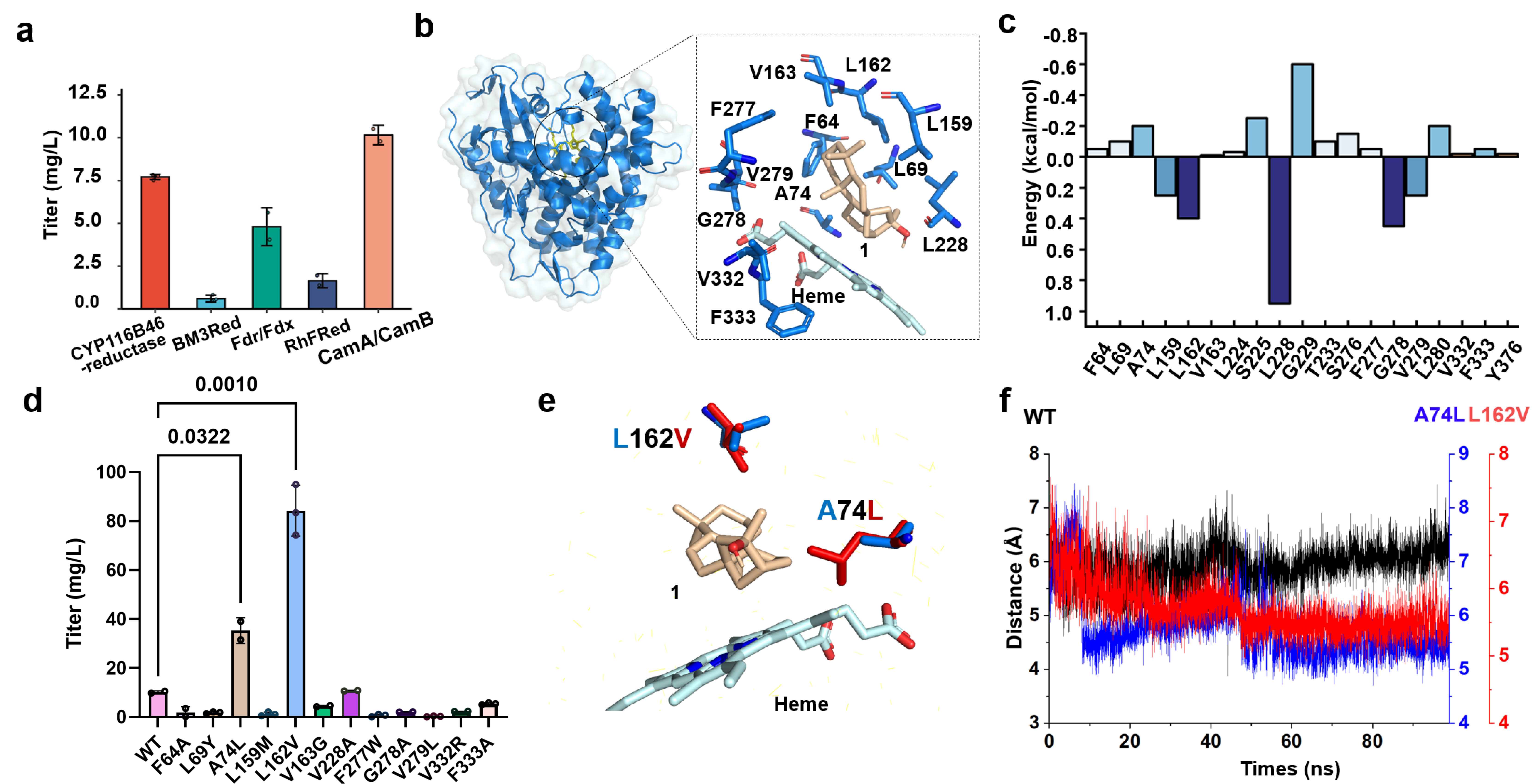
this work

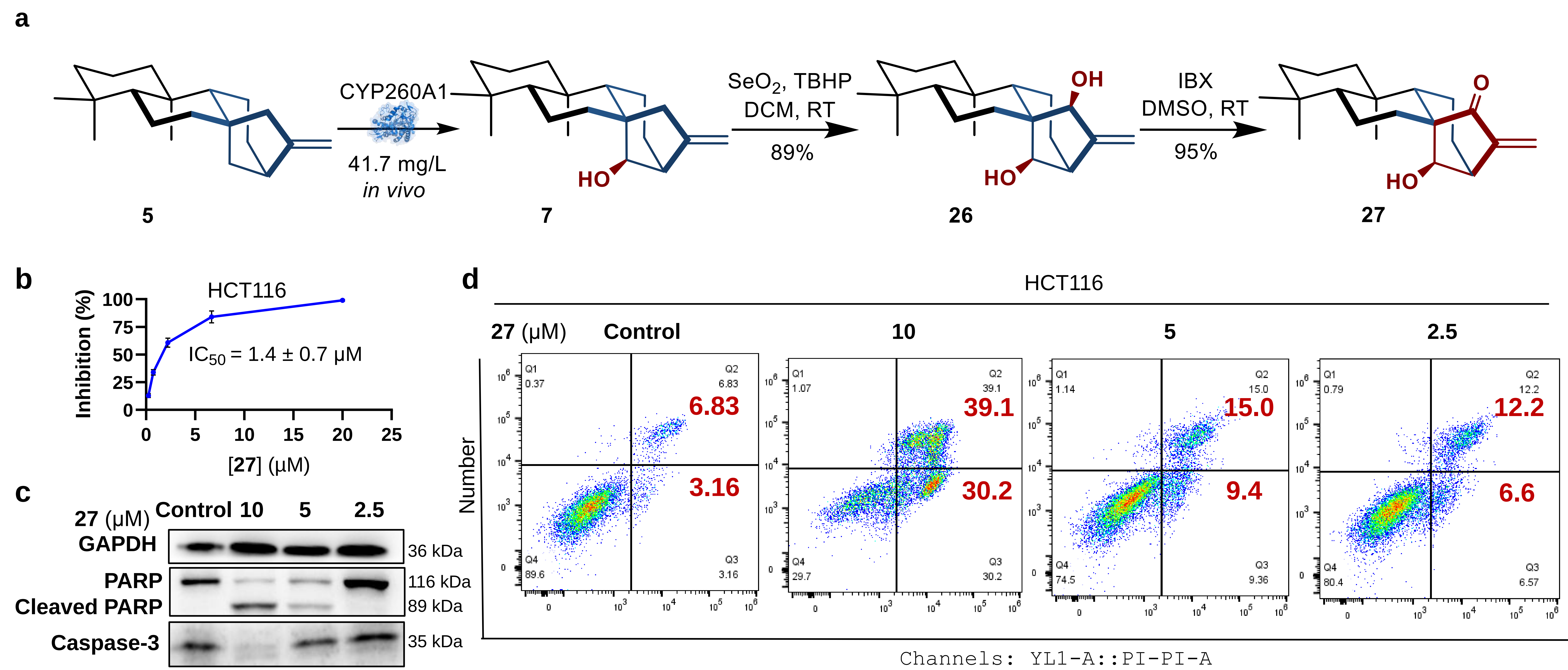
- CHS strategy identified three bacterial C-14 hydroxylases
- CYP260A1 engineering improved catalytic efficiency (84.2 mg/L)
- a potent cytotoxic **27**, $IC_{50} 1.4 \mu M$

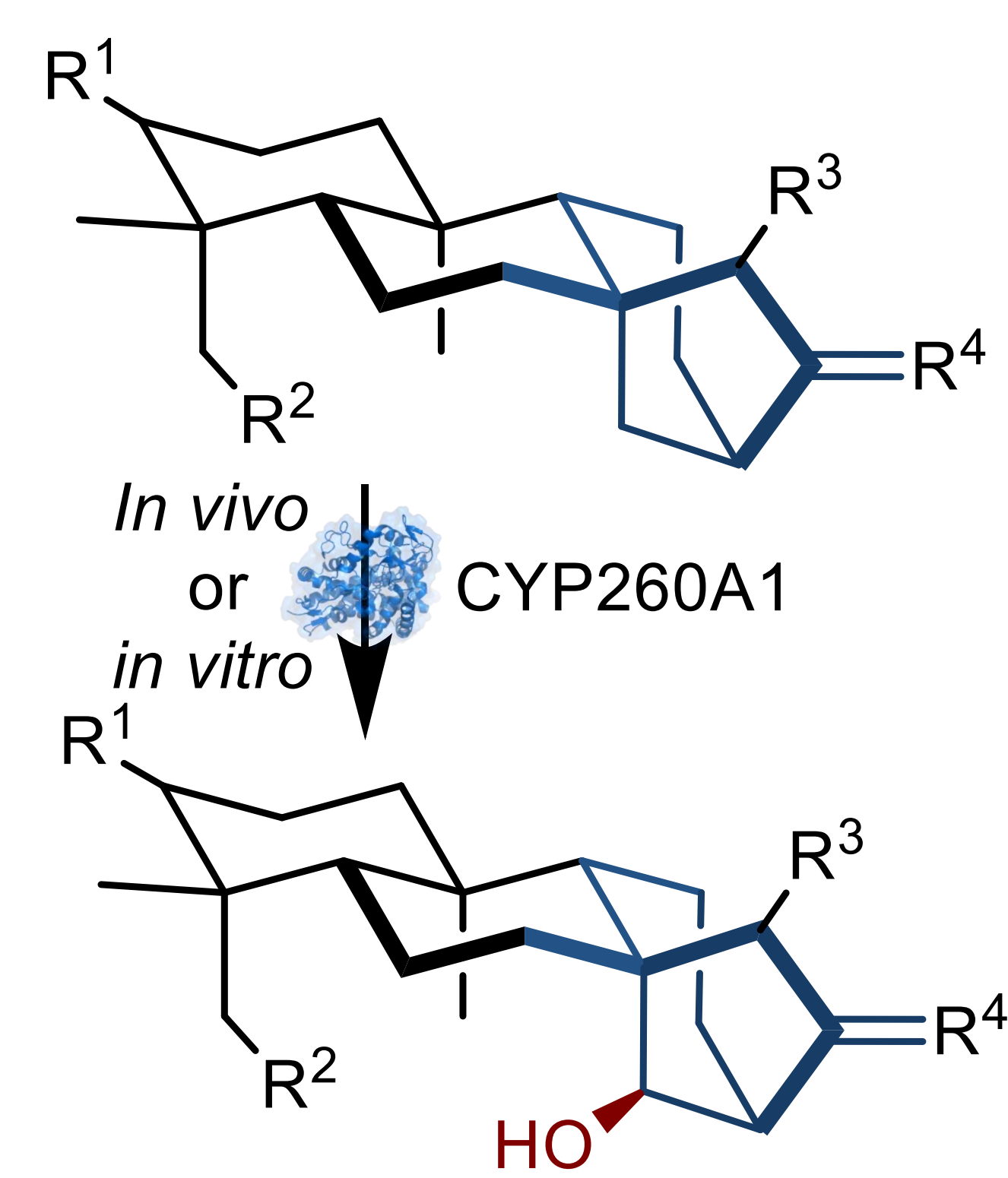
CHS strategy, enzyme engineering, SAR evaluation

a CHS strategy utilization**b Enzyme engineering****c SAR evaluation**









Substrate

1 : $R^1, R^2, R^3 = H, R^4 = \beta\text{-Me}, \alpha\text{-OH}$

5 : $R^1, R^2, R^3 = H, R^4 = CH_2$

6 : $R^1 = \beta\text{-OH}, R^2, R^3 = H, R^4 = CH_2$

8 : $R^1, R^2 = H, R^3 = \alpha\text{-OH}, R^4 = CH_2$

9 : $R^1, R^2 = H, R^3 = O, R^4 = CH_2$

12 : $R^1 = O, R^2, R^3 = H, R^4 = CH_2$

24 : $R^1 = H, R^2 = OH, R^3 = H, R^4 = CH_2$