Applied and Environmental Microbiology

Applied and Industrial Microbiology | Full-Length Text

AlmA involved in the long-chain *n*-alkane degradation pathway in *Acinetobacter baylyi* ADP1 is a Baeyer–Villiger monooxygenase

Chao-Fan Yin,[1] Yong Nie,[2] Tao Li,[1] Ning-Yi Zhou[1]

AUTHOR AFFILIATIONS See affiliation list on p. 12.

ABSTRACT Many *Acinetobacter* species can grow on *n*-alkanes of varying lengths (≤C40). AlmA, a unique flavoprotein in these *Acinetobacter* strains, is the only enzyme proven to be required for the degradation of long-chain (LC) *n*-alkanes, including C32 and C36 alkanes. Although it is commonly presumed to be a terminal hydroxylase, its role in *n*-alkane degradation remains elusive. In this study, we conducted physiological, biochemical, and bioinformatics analyses of AlmA to determine its role in *n*-alkane degradation by *Acinetobacter baylyi* ADP1. Consistent with previous reports, gene deletion analysis showed that *almA* was vital for the degradation of LC *n*-alkanes (C26–C36). Additionally, enzymatic analysis revealed that AlmA catalyzed the conversion of aliphatic 2-ketones (C10–C16) to their corresponding esters, but it did not conduct *n*-alkane hydroxylation under the same conditions, thus suggesting that AlmA in strain ADP1 possesses Baeyer–Villiger monooxygenase (BVMO) activity. These results were further confirmed by bioinformatics analysis, which revealed that AlmA was closer to functionally identified BVMOs than to hydroxylases. Altogether, the results of our study suggest that LC *n*-alkane degradation by strain ADP1 possibly follows a novel subterminal oxidation pathway that is distinct from the terminal oxidation pathway followed for short-chain *n*-alkane degradation. Furthermore, our findings suggest that AlmA catalyzes the third reaction in the LC *n*-alkane degradation pathway.

IMPORTANCE Many microbial studies on *n*-alkane degradation are focused on the genes involved in short-chain *n*-alkane (≤C16) degradation; however, reports on the genes involved in long-chain (LC) *n*-alkane (>C20) degradation are limited. Thus far, only AlmA has been reported to be involved in LC *n*-alkane degradation by *Acinetobacter* spp.; however, its role in the *n*-alkane degradation pathway remains elusive. In this study, we conducted a detailed characterization of AlmA in *A. baylyi* ADP1 and found that AlmA exhibits Baeyer–Villiger monooxygenase activity, thus indicating the presence of a novel LC *n*-alkane biodegradation mechanism in strain ADP1.

KEYWORDS *Acinetobacter baylyi* ADP1, AlmA, Baeyer–Villiger monooxygenase (BVMO), long-chain (LC) *n*-alkane, biodegradation

Long-chain (LC) *n*-alkanes are organic compounds with >20 carbon atoms in their skeleton. They form the solid and inert component of crude oil, which when spilled can have severe impacts on land and marine environments (1). Although LC *n*-alkanes are valuable inert oil reserves, their recovery and transport can be challenging (2). Microbial-enhanced remediation and oil recovery of oil-contaminated sites have thus gained increasing interest (3, 4). Thus far, several *n*-alkane-degrading bacterial strains have been identified (5, 6), including *Acinetobacter* spp., which utilize LC *n*-alkanes. Notably, *Acinetobacter baylyi* ADP1, which is capable of degrading LC *n*-alkanes (≤C36) (5, 7), has been identified as a naturally competent organism, making it suitable for genetic manipulation (8, 9).

Editor Haruyuki Atomi, Kyoto University, Katsura, Nishikyo-ku, Kyoto, Japan

Address correspondence to Ning-Yi Zhou, ningyi.zhou@sjtu.edu.cn.

The authors declare no conflict of interest.

See the funding table on p. 13.

Received 14 September 2023
Accepted 22 November 2023
Published 3 January 2024

Two biodegradation pathways, namely, terminal and subterminal oxidation pathways, have been proposed for *n*-alkane degradation (10–12) (Fig. 1). The terminal oxidation pathway has been extensively studied (13–15); however, the subterminal oxidation pathway and its components remain largely unexplored (16). Moreover, various short-chain (SC) *n*-alkane (<C16) degrading enzymes, especially terminal hydroxylases, such as cytochrome P450 (17), membrane-spanning metalloenzyme AlkB (18, 19), and AlkM (7, 20), have been well-studied. In contrast, only two LC *n*-alkane degrading enzymes, namely, LadA and AlmA, have been reported in the literature. LadA is a flavin mononucleotide-dependent two-component flavoprotein obtained from *Geobacillus thermodenitrificans* NG80-2 (21), and it has been biochemically identified as an LC *n*-alkane terminal hydroxylase. AlmA is a flavin adenine dinucleotide (FAD)-dependent single-component flavoprotein obtained from *Acinetobacter* spp., including strains DSM 17874 and ADP1 (22), and it has been demonstrated to be involved in LC *n*-alkane (C32 and C36) degradation. *In vitro* studies found that an AlmA homolog from *Alcanivorax dieselolei* B-5 could catalyze the terminal hydroxylation of C10–C36 alkanes (13), while an AlmA homolog from *Acinetobacter radioresistens* S13 could catalyze Baeyer–Villiger oxidation of 4-phenyl-2-butanone (a non-alkane metabolite) and *S*-oxidation of

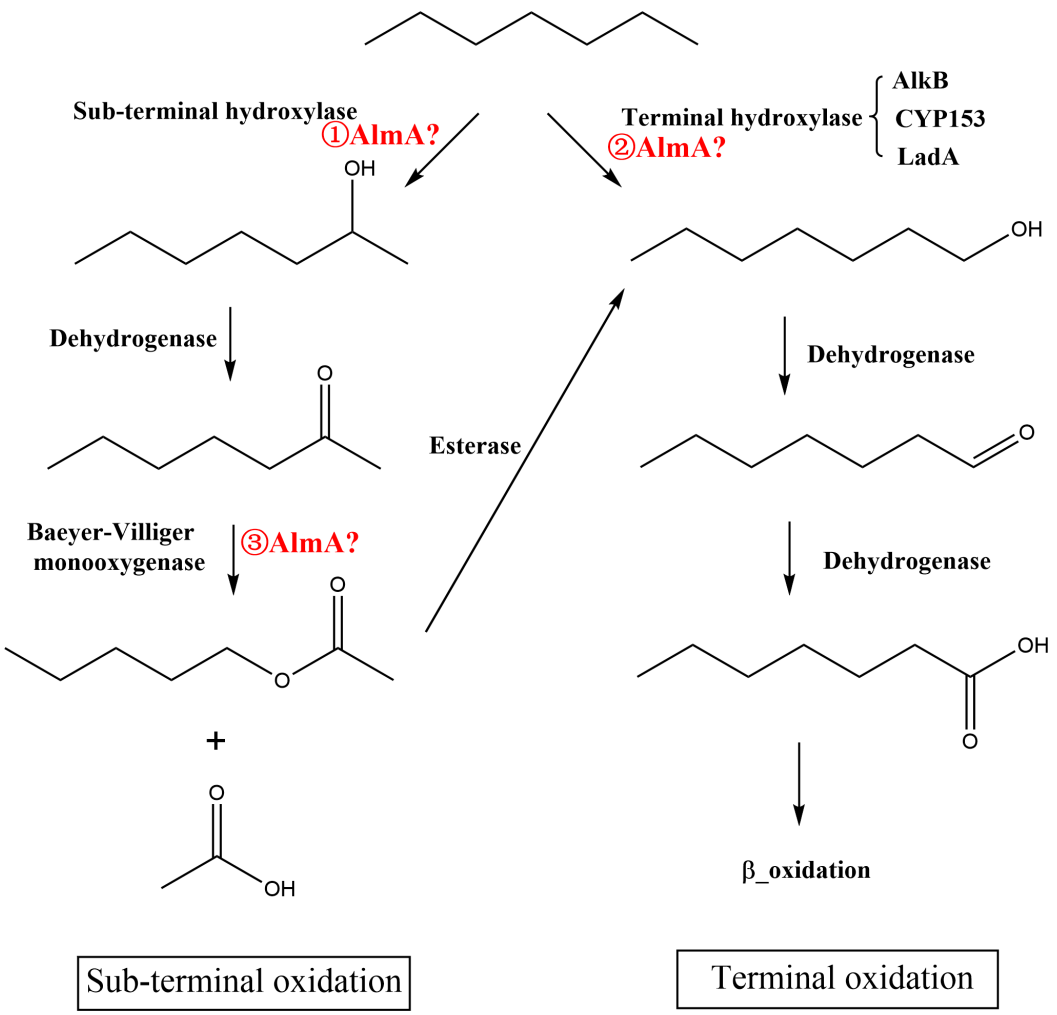


FIG 1 The terminal and subterminal oxidation pathways for aerobic *n*-alkane degradation. Terminal oxidation pathway and its terminal hydroxylase systems, including AlkB (18), CYP153 (17), and LadA (11, 21), have been extensively studied. AlmA, a key oxygenase, is proposed to play one of the three putative roles in the terminal or subterminal pathways that are highlighted in red.

ethionamide (23). These differing reports on AlmA activity make it difficult to discern the exact role of AlmA in LC *n*-alkane degradation by strain ADP1.

In this study, we conducted physiological, biochemical, and bioinformatics analyses to identify the functional role of AlmA in LC *n*-alkane degradation by *A. baylyi* ADP1.

RESULTS

Gene *almA* is needed for LC *n*-alkane (>C26) degradation by *A. baylyi* ADP1

To analyze the physiological function of *almA* in *n*-alkane degradation by strain ADP1, we constructed an *almA*-deleted strain ADP1Δ*almA* and conducted growth assays using *n*-alkanes of different chain lengths as the sole carbon source. The mutant strain grew normally on C24 and grew poorly on C26 and C28, but did not grow on C30, C32, and C36 (Fig. 2). These results indicate that AlmA functions specifically in LC *n*-alkane (≥C26) utilization in strain ADP1, consistent with the previous reports on *almA* function *in Acinetobacter* sp. strain DSM 17874 (22) and *A. dieselolei* B-5 (13).

AlmA was presumed to be a hydroxylase catalyzing the terminal hydroxylation of LC *n*-alkanes to produce their corresponding primary alcohols (5). Therefore, theoretically, the deletion of *almA* gene would not affect its utilization of the corresponding aliphatic primary alcohols. To verify this hypothesis, the corresponding aliphatic primary alcohols, C32-OH, C30-OH, C28-OH, C26-OH, and C24-OH, were used as the sole carbon sources for growth assays of the strain ADP1Δ*almA*. The results revealed that the mutant strain did not grow on C32-OH, grew poorly on C30-OH and C28-OH, but grew normally on C26-OH and C24-OH (Fig. S1). However, complementation with $almA_{tru5aa}$ (a truncated but functional *almA* that could be readily purified from *Escherichia coli* as described below) restored the ability of the mutant strain to grow on C32-OH, C30-OH, and C28-OH to virtually the wild-type levels (Fig. S1). These results suggest that instead of terminal hydroxylation, AlmA possesses other uncharacterized functions in the LC *n*-alkane

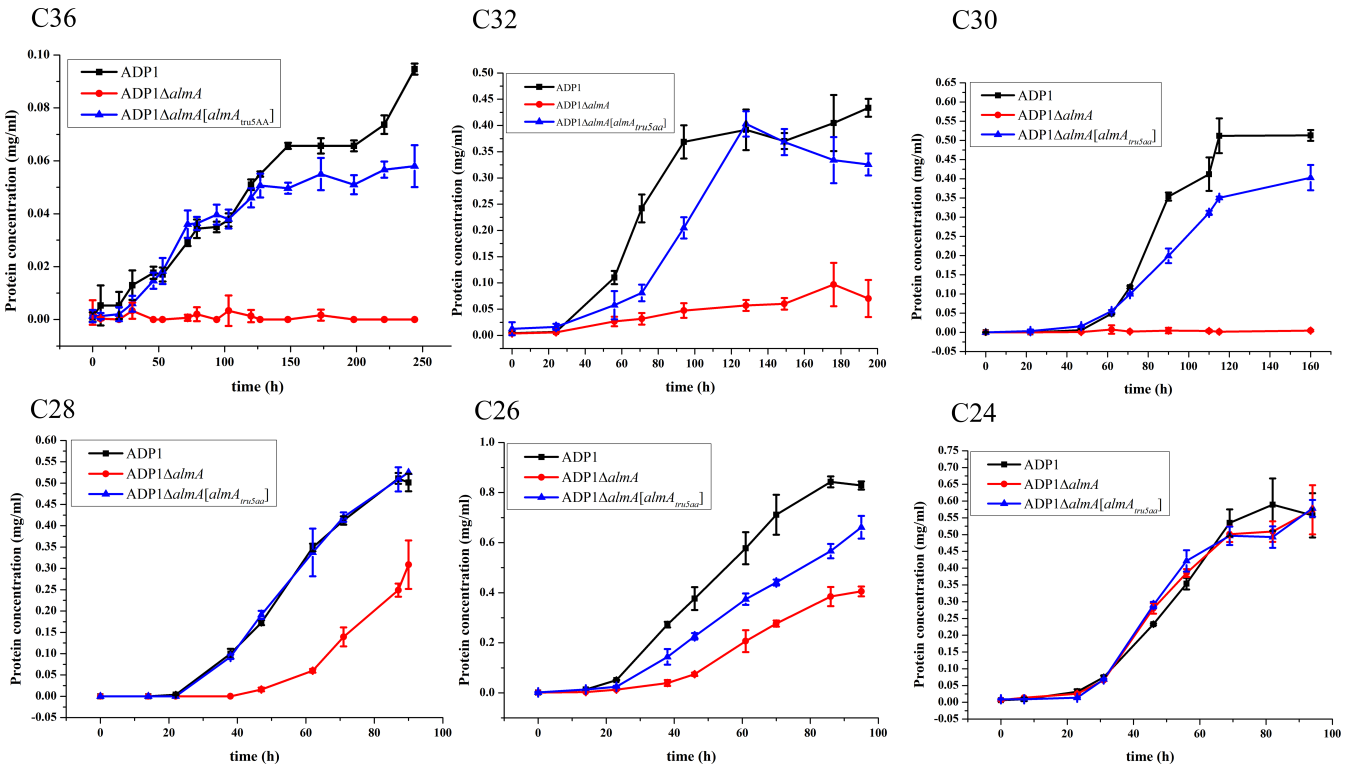


FIG 2 Growth of wild-type, *almA*-deleted, and $almA_{tru5aa}$-complemented strains on various long-chain *n*-alkanes. Strains were grown in mineral salts medium supplemented with various *n*-alkanes (1%, wt/vol). Growth was measured as an increase in protein content over time. The values represent mean ± SD of three independent experiments.

degradation pathway. However, further growth assays were needed using aliphatic LC secondary alcohols to discern the physiologic role of AlmA in LC *n*-alkane degradation by strain ADP1.

To determine if AlmA is the enzyme for the first reaction in the LC *n*-alkane degradation pathway, we conducted removal assays of C32 by the wild-type and ADP1Δ*almA* strain (Fig. S2). The results revealed that both the wild-type and mutant ADP1Δ*almA* strains could consume C32 in a similar manner. However, no transformation products were detected in this experiment. These results indicate the presence of uncharacterized enzymes that catalyze the first step in the LC *n*-alkane degradation pathway in the strain ADP1.

AlmA contains the Baeyer–Villiger monooxygenases (BVMOs) motif and clustered with BVMOs

Baeyer–Villiger monooxygenases (BVMOs) form a distinct class of flavoproteins that catalyze the insertion of an oxygen atom in a C-C bond using dioxygen and NAD(P)H (24). Although BVMOs and flavin-containing hydroxylases catalyze different reactions, they share a similar electron transport system and catalytic mechanism, and because of that, they share certain sequence similarities (25). This may lead to difficulties in distinguishing between the two enzymes.

To analyze the putative physiological and biochemical functions of AlmA, we constructed a maximum likelihood (ML) phylogenetic tree and conducted multiple sequence alignment (MSA) using functionally identified homologous proteins of AlmA on the basis of their amino acid sequences. As shown in Fig. 3a, AlmA was closely clustered with functionally identified BVMOs and distinctly separated from hydroxylases, thus indicating that AlmA most likely functions as a BVMO rather than a hydroxylase. This was further confirmed by the results of MSA (Fig. 3b). AlmA contains all of the conserved residues and the unique motif (FxGxxxHxxxW(P/D)) of the BVMOs (25). The strictly conserved residues of AlmA are shown in its predicted 3D structure in Fig. 3c. The conserved AlmA residues were located in the FAD- and NADPH-binding sites, indicating that AlmA, as a BVMO, shared conserved but distinct cofactor-binding mechanisms with hydroxylases, which may in turn determine the reaction types, such as BV oxidation or hydroxylation (26). However, further crystallographic analysis of AlmA is required to elucidate the exact mechanism underlying this phenomenon.

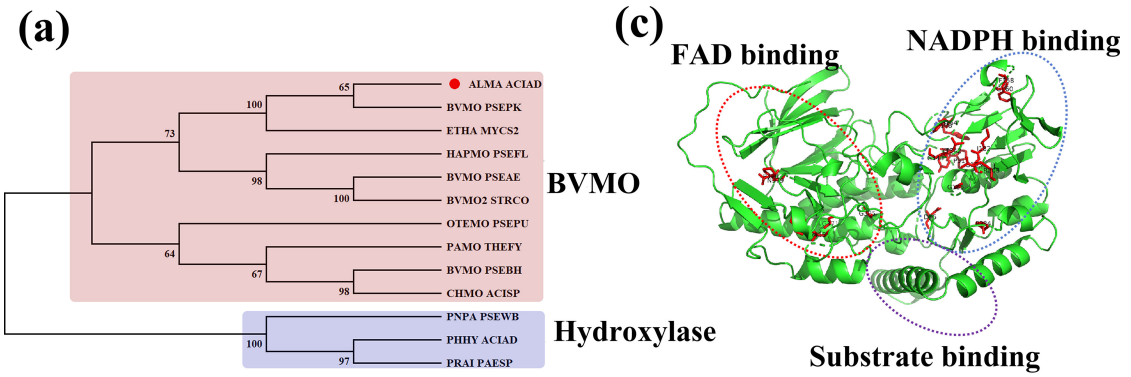
$AlmA_{tru5aa}$ functions as a BVMO

Heterologous expression of AlmA led to the formation of inclusion bodies in *E. coli*; however, heterologous expression of $AlmA_{tru5aa}$ (with N-terminal 5aa deletion) in *E. coli* was effective and the protein was successfully purified (Fig. S3). To verify if $AlmA_{tru5aa}$ retained the wild-type function of AlmA, it was complemented into an *almA*-deletion mutant strain, ADP1Δ*almA*, to yield strain ADP1Δ*almA*[$almA_{tru5aa}$]. The complemented strain was observed to show a similar growth pattern as the wild-type strain (Fig. 2), indicating that $AlmA_{tru5aa}$ retained the full physiological function in strain ADP1.

The enzyme activity of the recombinant His_6-$AlmA_{tru5aa}$ was then assayed using *n*-alkanes and aliphatic ketones listed in Table S1. The results revealed that His_6-$AlmA_{tru5aa}$ catalyzed the conversion of the aliphatic 2-ketones, C10-2-one, C12-2-one (Fig. 4a), C13-2-one, C14-2-one, and C16-2-one (Fig. 4b) to their corresponding alkyl esters (alkanol-acetates), whose identities were confirmed by gas chromatography–mass spectrometry (GC–MS; Fig. 4). In contrast, no hydroxylase activity was detected for His_6-$AlmA_{tru5aa}$ toward *n*-alkanes under the same conditions. These results indicate that AlmA functions as a BVMO in the LC *n*-alkane degradation pathway.

Molecular docking simulations of AlmA

Protein docking was conducted using 2-hexadecanone (C16-2-one) to model potential interactions. As shown in Fig. 5a, a large double-entry cavity (4,616 $Å^3$) leading to the



(b)

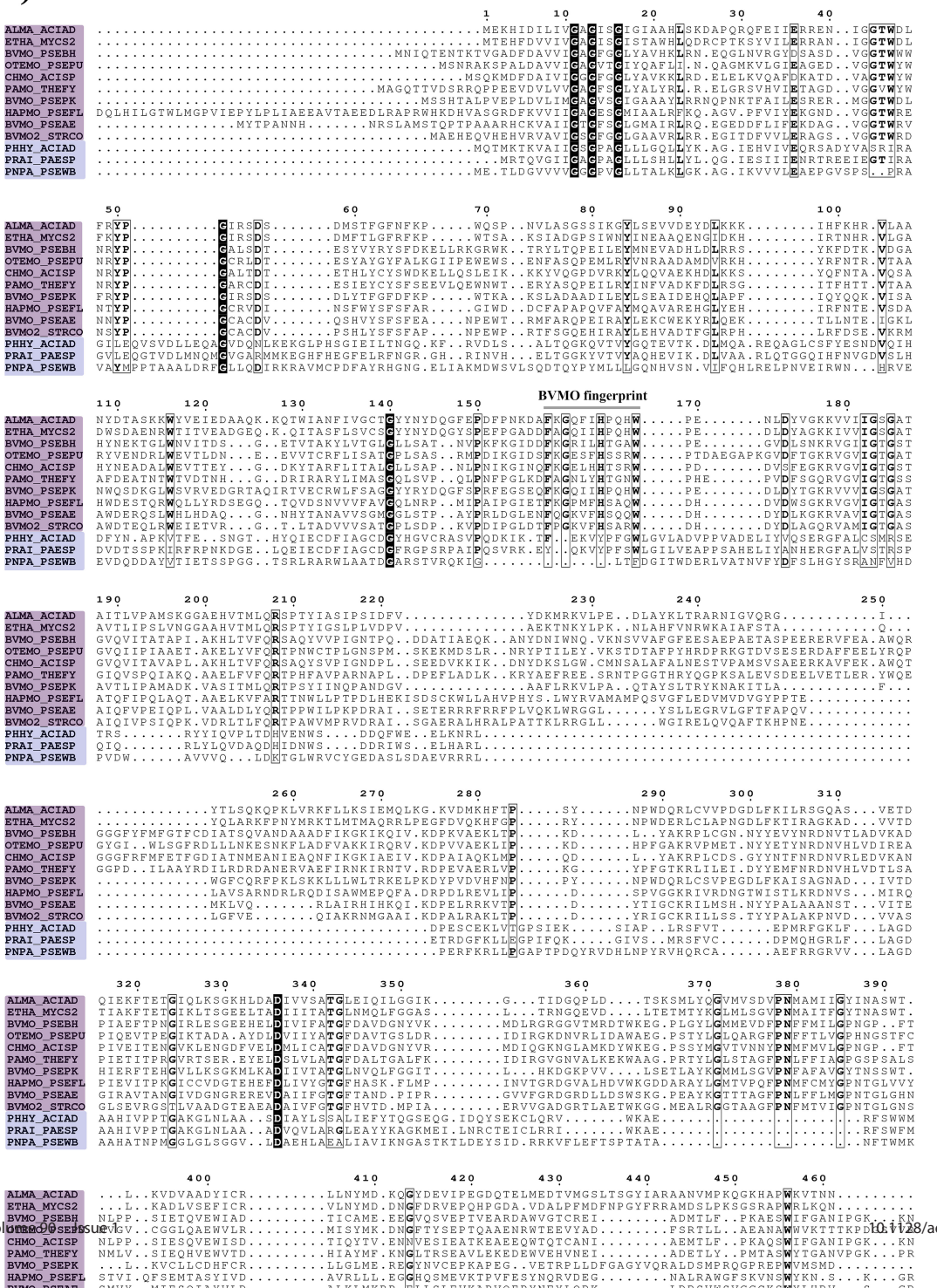


FIG 3 (Continued)

FIG 3 Sequence analysis of AlmA. (a) The phylogenetic tree of AlmA and its homologs. The tree was constructed with 1,000 bootstrap replicates using the maximum likelihood method in MEGA v11 (27). The AlmA from *Acinetobacter baylyi* ADP1 is marked with a red dot. The clades of functionally identified Baeyer–Villiger monooxygenases (BVMOs) and hydroxylases are indicated with red and blue backgrounds, respectively. (b) Multiple sequence alignment of AlmA and its homologs.The sequences used for the construction of the phylogenetic tree were aligned by Clustal Omega (28) and rendered with ESPript (29). The black boxes highlight the relatively conserved residues, and the black-filled boxes highlight the absolutely conserved residues. The BVMO fingerprint is indicated with a black line. (c) The AlphaFold2-predicted 3D structure of AlmA. The conserved residues identified by MSA analysis are presented as red sticks. The predicted cofactor- and substrate-binding domains are indicated by dashed ovals.

active sites was observed in AlmA, which may be an entry point for LC ketones and NADPH into the AlmA active sites. Additionally, the results revealed that certain amino acid (aa) residues, including Ile250, Val261, Phe264, Leu265, and Trp289, formed a hydrophobic environment (Fig. 5b; Fig. S4) in the active sites, possibly to accommodate the hydrophobic aliphatic chains of LC ketones. Additionally, the aa residues Asp56 and Arg292, adjacent to FAD, probably aided in the proper positioning of LC ketones to enable hydrogen bond formation between its unique carbonyl oxygen and the cofactors (30).

To confirm this hypothesis, we constructed two mutants of $AlmA_{tru5aa}$, D56A, and R292A, and then conducted enzyme activity assays using GC–MS (Fig. 6a). Compared to $AlmA_{tru5aa}$, $AlmA_{tru5aa}$-D56A and $AlmA_{tru5aa}$-R292A lost their ability to convert C12-2-one to its corresponding ester, indicating that Asp56 and Arg292 are essential for the BVMO activity of $AlmA_{tru5aa}$. To further confirm the role of AlmA as a BVMO in LC-chain *n*-alkane degradation, the D56A- and R292A-encoding genes were complemented into strain ADP1Δ*almA* to form ADP1Δ*almA*[$almA_{tru5aa}$-*D56A*] and ADP1Δ*almA*[$almA_{tru5aa}$-*R292A*]. Their growth on C32 revealed that the $AlmA_{tru5aa}$-D56A and $AlmA_{tru5aa}$-R292A mutant strains were unable to restore the ability of the strain ADP1Δ*almA* to grow on C32 (Fig. 6b), unlike the $almA_{tru5aa}$ mutant. These results indicate that AlmA functions as a BVMO in LC *n*-alkane degradation, which further suggests the presence of a subterminal LC *n*-alkane degradation pathway in strain ADP1.

DISCUSSION

Several studies found that *almA* is essential for LC *n*-alkane degradation by *Acinetobacter* spp. (22). It was largely presumed that the flavoprotein AlmA was a hydroxylase and that LC *n*-alkane degradation followed the terminal oxidation pathway similar to SC *n*-alkane degradation (5). However, in the present study, AlmA was identified as a BVMO, suggesting that LC *n*-alkane degradation by *Acinetobacter* spp. possibly follows a distinct subterminal oxidation pathway instead of the terminal oxidation pathway. Compared to the terminal oxidation pathway, the subterminal oxidation pathway is more efficient for carbon acquisition during the LC *n*-alkane degradation into SC alcohols (31). Therefore, bacterial strains may adopt the subterminal oxidation pathway to thrive on complex substrates, such as LC *n*-alkanes. However, further prospective studies, including structural analysis of AlmA, enzymatic assays using LC aliphatic ketones (such as C32-2-one), and identification of the initial subterminal hydroxylase for LC *n*-alkanes degradation, are required to verify this hypothesis.

The genetic organization of the *almA* gene cluster in strain ADP1 is highly similar to that of its counterparts in the other strains of *Acinetobacter* spp., but different from the *almA* gene clusters in other genera, such as *Alcanivorax dieselolei* B-5 (32) (Fig. S5). These results indicate that *n*-alkane degrading enzymes in different genera probably have unique evolutionary origins and functions, despite sharing sequence similarity.

There are various bacterial enzymes for the activation and degradation of inert *n*-alkanes of varied chain lengths. AlmA (22), AlkB (18, 19), LadA (11, 21), and CYP153 (17) are four signature *n*-alkane oxygenases that are commonly found in different bacterial genera. Among these, AlkB, LadA, and CYP153 have been established as *n*-alkane terminal hydroxylases (5), while AlmA was presumed to be a terminal hydroxylase. However, the findings of our study revealed that AlmA functions as a BVMO and that

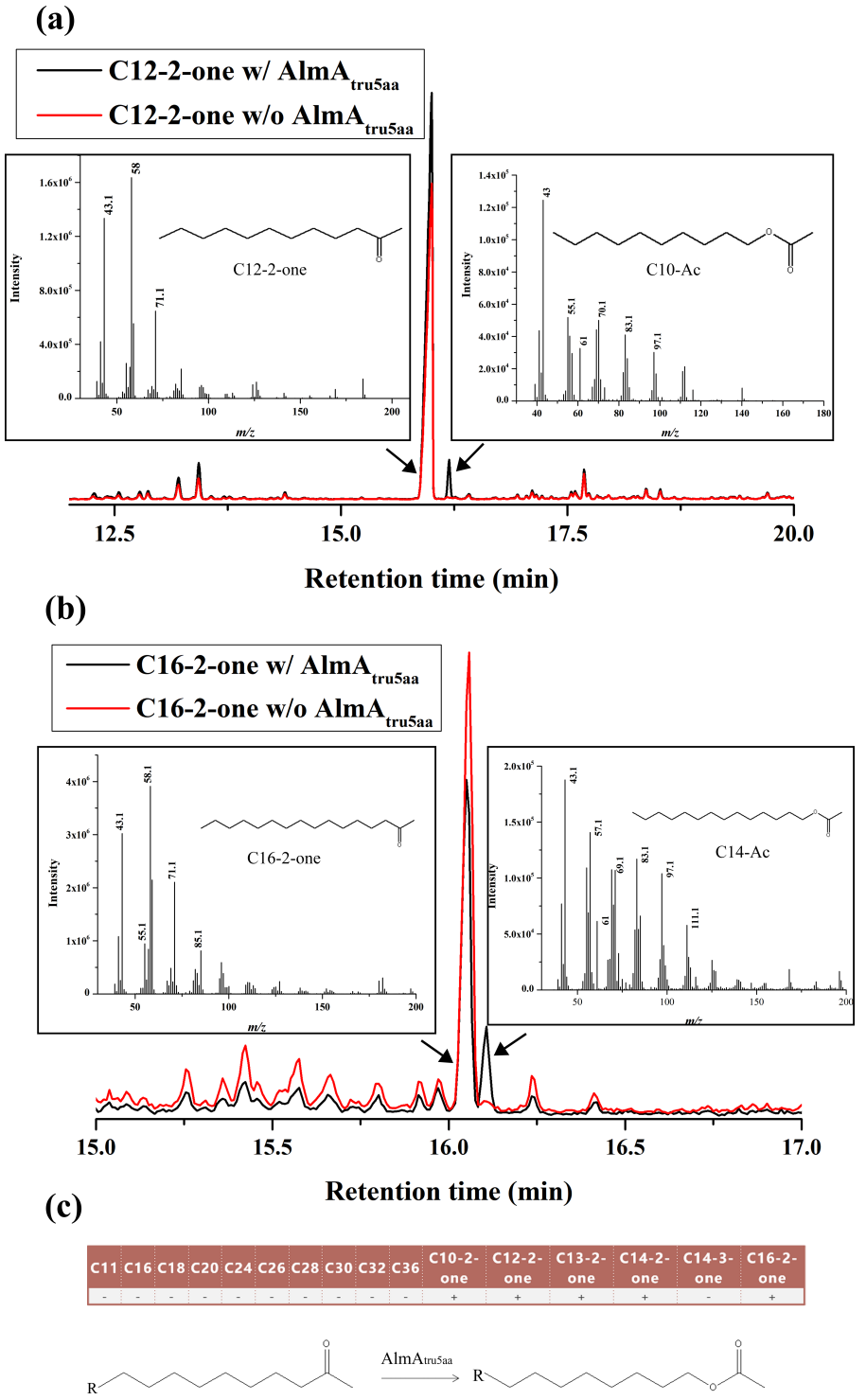


FIG 4 Enzyme activity assay of $AlmA_{tru5aa}$ and product identification using gas chromatography–mass spectrometry (GC–MS). GC traces show that the 2-dodecanone (C12-2-one) (a) and 2-hexadecanone (C16-2one) (b) were converted to their corresponding aliphatic esters by purified $AlmA_{tru5aa}$. The boxes adjacent to GC traces show the MS spectra of the substrates and their corresponding products. The red trace shows the negative control without the enzyme treatment. (c) The table lists the substrates used in this study, and the signs "+" and "–" represent the activity and inactivity of $AlmA_{tru5aa}$, respectively, against the corresponding substrates. The figure displays the Baeyer–Villiger reaction catalyzed by $AlmA_{tru5aa}$ on aliphatic ketones.

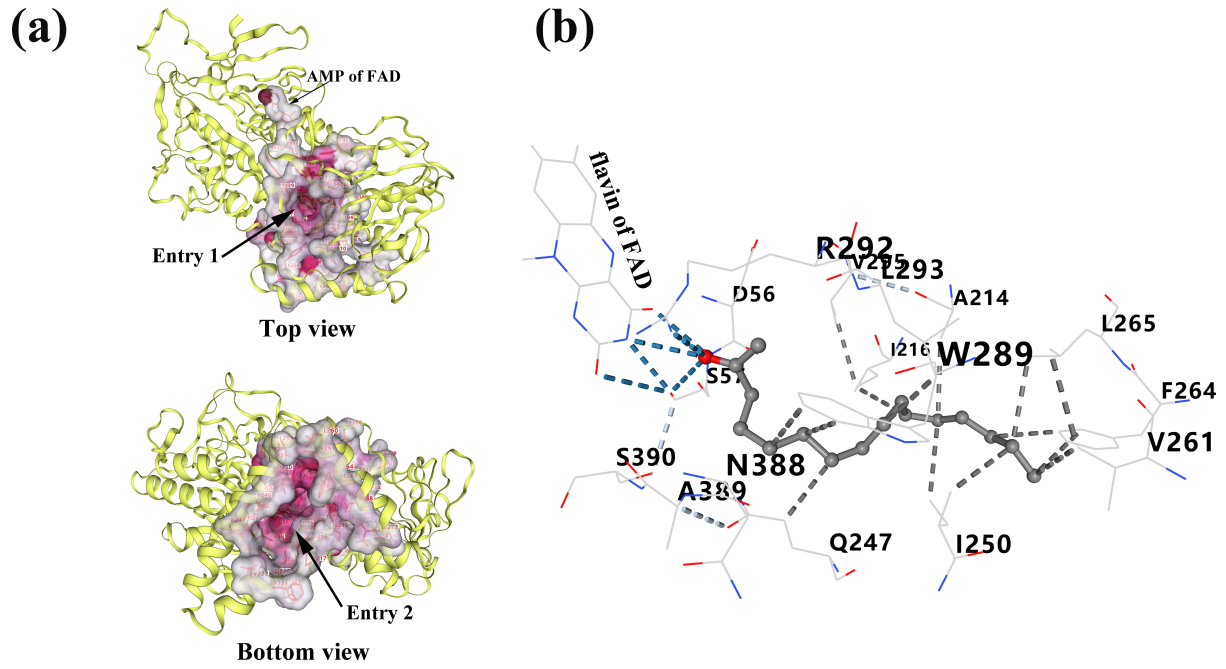


FIG 5 Molecular docking models of 2-hexadecanone (C16-2-one) in AlmA. (a) The putative substrate-binding cavity of AlmA. The predicted 3D structure of AlmA is shown as yellow cartoon, the detected putative substrate-binding cavity (including entry1 and entry2) is shown as rendered-surface, and the binding-related residues are shown as red sticks. (b) The putative interaction model of AlmA and C16-2-one. Docked C16-2-one is shown as a ball-and-stick model and the cofactor FAD and key residues are shown as sticks. Hydrogen bonds are represented by blue dotted lines, and hydrophobic interactions are represented by gray dotted lines.

it may be involved in the subterminal oxidation of LC *n*-alkanes. Although the physiological substrate range of these four oxygenases has yet to be fully established, the coexistence and duplication of their encoding genes have been observed in certain bacterial strains (16, 32, 33), indicating that bacterial strains can extend their *n*-alkane degradation range via gene redundancy and duplication. Distribution analysis of the genes encoding

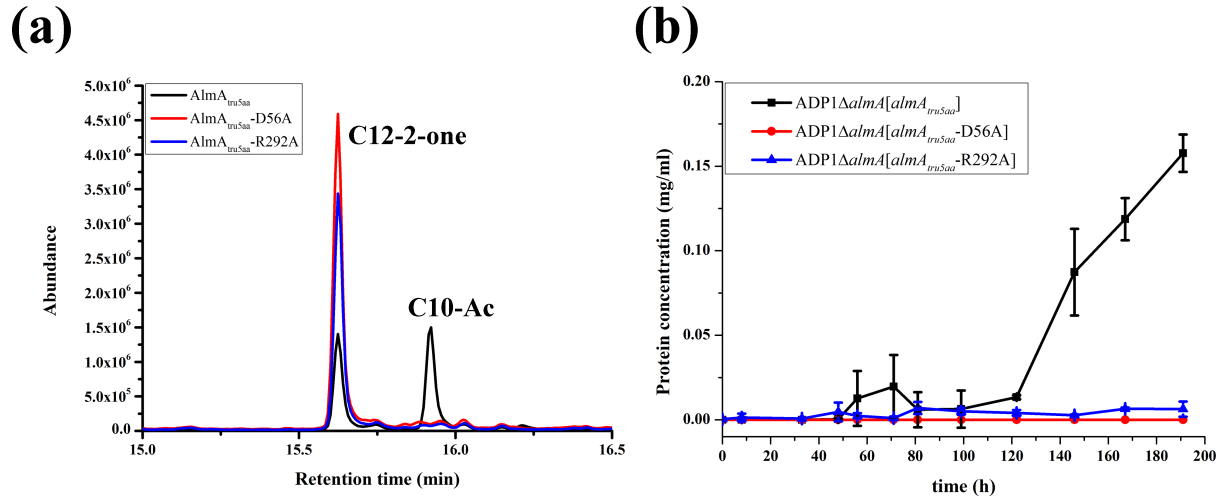


FIG 6 Biochemical and physiological identification of the AlmA$_{tru5aa}$ mutants, D56A and R292A. (a) Gas chromatography traces show that purified AlmA$_{tru5aa}$ converted C12-2-one to its corresponding ester (c10-ac), while AlmA$_{tru5aa}$-D56A and AlmA$_{tru5aa}$-R292A (indicated by red and blue, respectively) were unable to catalyze this reaction. (b) Growth of ADP1Δ*almA* strains complemented with *almA*$_{tru5aa}$, *almA*$_{tru5aa}$*-D56A*, and *almA*$_{tru5aa}$*-R292A* on C32. Complementation with *almA*$_{tru5aa}$ restored the growth of the strain ADP1Δ*almA* on C32; however, complementation with *almA*$_{tru5aa}$*-D56A* and *almA*$_{tru5aa}$*-R292A* did not restore its growth on C32. .Growth was measured as an increase in protein content over time. The values presented are the averages of three independent experiments, and the error bars indicate standard deviation.

these signature *n*-alkane oxygenases in bacterial genomes might give some insights into this hypothesis. Therefore, we conducted an overall query using these four signature enzymes against all of the available bacterial genomes (up to May 2023; *N* = 32,796 genomes) with whole-genome assembly. As shown in Fig. 7a, LadA, AlmA, AlkB, and CYP153 were found in 6,191, 3,761, 2,957, and 619 bacterial genomes, respectively, suggesting their ubiquitous distribution in bacterial genomes.

In addition, the coexistence of these four enzymes has also been observed in some bacterial species. Among the bacteria (*N* = 7,788) containing one or more of the

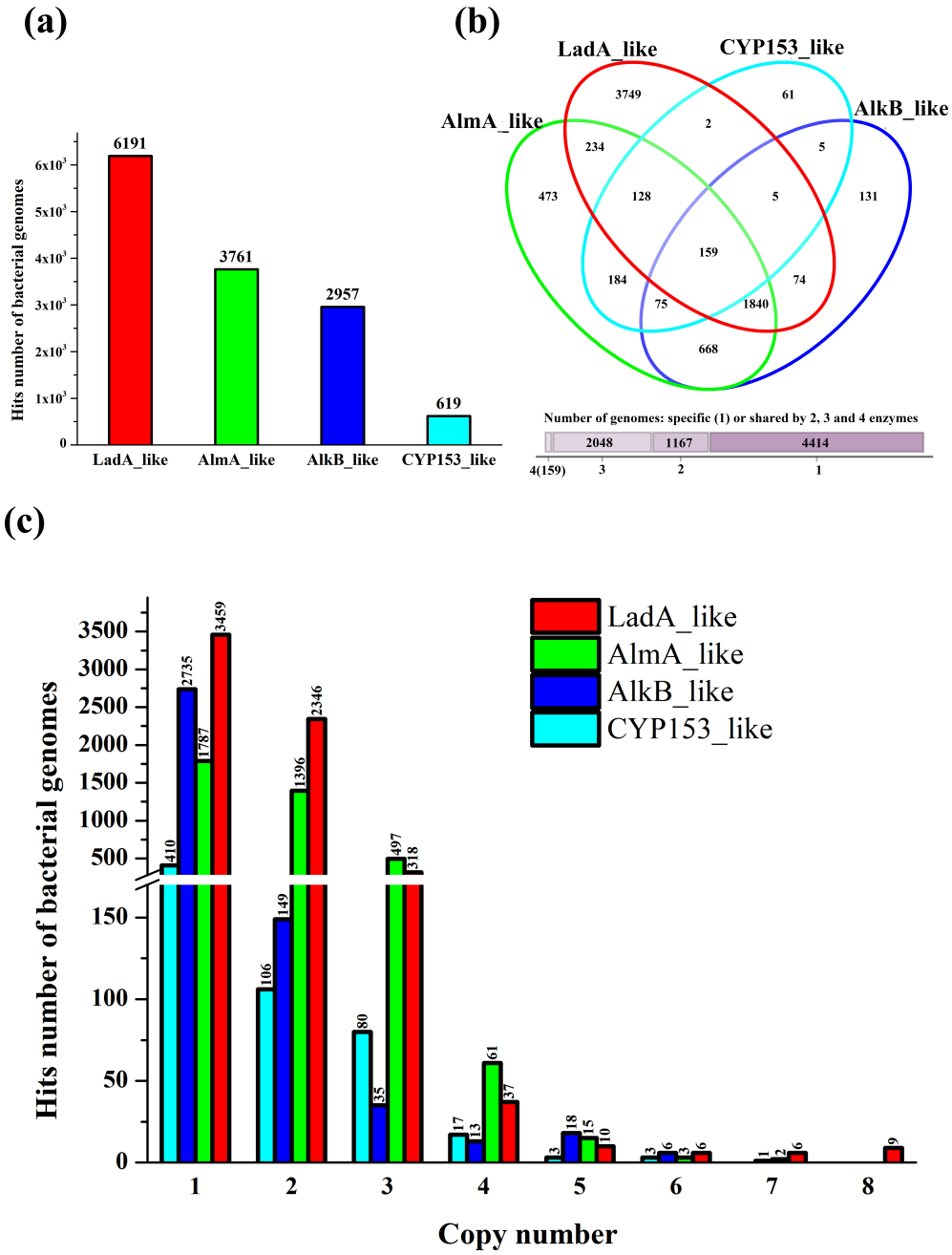


FIG 7 Distribution and coexistence of the genes encoding signature *n*-alkane oxygenases in bacterial genomes. (a) Histogram shows the number of bacterial whole-genomes containing the signature *n*-alkane oxygenases, LadA, AlmA, AlkB, and CYP153. (b) Venn diagram shows the coexistence of the four *n*-alkane oxygenases in the bacterial genomes. (c) Histogram shows the copy number of the genes encoding *n*-alkane oxygenases in bacterial genomes. The genes encoding LadA, AlmA, AlkB, and CYP153 are represented by red, green, blue, and cyan, respectively.

signature enzymes, 43% contained at least two of the four enzymes and 28% contained at least three of the four enzymes (Fig. 7b). Additionally, AlkB and AlmA were the most commonly coexisting enzymes, while LadA was commonly found alone in the bacterial genomes, suggesting its unique genetic origin.

Moreover, we observed the presence of multiple gene copies of these signature *n*-alkane oxygenases in the bacterial genomes (Fig. 7c), which were previously suggested to be used for degrading *n*-alkanes of varying chain lengths (33). The broad distribution, coexistence, and duplication of the genes encoding signature *n*-alkane oxygenases indicate that the *n*-alkane-degrading bacteria probably share a common but complex catabolic network.

In this study, we conducted detailed physiological, biochemical, and bioinformatics analyses of AlmA from *A. baylyi* ADP1. The results revealed that AlmA exhibits BVMO activity and that LC *n*-alkane degradation possibly follows a subterminal degradation pathway that is distinct from the terminal degradation pathway used for SC *n*-alkane degradation.

MATERIALS AND METHODS

Chemicals, media, bacterial strains, plasmids, and primers

All of the bacterial strains, plasmids, and primers used in this study are listed in Table 1. All of the *n*-alkanes, aliphatic primary alcohols, and aliphatic ketones used in this study (Table S1) were purchased from Sigma (St. Louis, MO, USA) and Aladdin Reagent Co. Ltd. (Shanghai, China). *A. baylyi* ADP1 (9) and its mutants were cultured in mineral salts medium (34) supplemented with various carbon sources (1%, wt/vol; as required) at 30°C and 180 rpm. Their growth was monitored by measuring the protein concentration in the cultures using the BCA Protein Assay Kit (Beyotime, China). *E. coli* strains were grown in lysogeny broth (LB) (35) supplemented with 50 µg/mL kanamycin or 100 µg/mL ampicillin at 37°C and 180 rpm.

Gene deletion and complementation

Being naturally competent, strain ADP1 can be genetically manipulated by incubating its log-phase cultures with linear DNA constructs (8, 9). The deletion and complementation assays of *almA* (GenBank GeneID: 45235407) were conducted using a previously described method (8) with minor modifications. For *almA* deletion, the upstream and downstream fragments of the target gene (from the strain ADP1 genome) were amplified and fused with a kanamycin resistance gene amplified from plasmid pK18mobsacB (36) by overlap extension polymerase chain reaction (PCR). The constructed donor DNA was transformed into strain ADP1, as described previously (8). The ADP1Δ*almA* double-crossover recombinants were screened on LB plates containing 50 µg/mL kanamycin and further verified by sequencing. $AlmA_{tru5aa}$, a variant of AlmA lacking N-terminal 5aa residues, was purified from *E. coli* as described below, and its encoding gene $almA_{tru5aa}$, with a 15-bp deletion (1–15) at the 5′ end (adding start codon (ATG) by primers), was complemented into strain ADP1Δ*almA* to yield strain ADP1Δ*almA*[$almA_{tru5aa}$]. The donor DNA was constructed by fusing homology arms, $almA_{tru5aa}$, and an ampicillin resistance gene amplified from pETduet-1 plasmid (Novagen). The strain ADP1Δ*almA*[$almA_{tru5aa}$] was confirmed by PCR amplification and sequencing, and $almA_{tru5aa}$ was expressed from its native promoter and Shine-Dalgarno sequences. Furthermore, the genes encoding $AlmA_{tru5aa}$ mutants, D56A and R292A, were complemented into strain ADP1Δ*almA* using the same method.

Protein expression and purification

The $almA_{tru5aa}$ gene was amplified and fused into pET-28a (+) (digested with NdeI and BamHI) using a One Step cloning kit (Vazyme Biotech Co., Ltd.). The resultant plasmids pET-$almA_{tru5aa}$ were transformed into *E. coli* BL21 (DE3) using a standard

TABLE 1 Bacterial strains, plasmids, and primers

Materials	Descriptions or sequences (5′→3′)	Sources or references
Bacterial strains		
E. coli DH5a	*lacU169ΔlacU169* (*φ80dlacZΔM15*) *hsdR17 recA1 endA1 hsdR17 thi*$^{-1}$ *gyrA96 relA1*	Novagen
E. coli BL21 (DE3)	F$^-$ *ompT hsdS*$_B$ (*Rb*$^-$*mB*$^-$) *gal* (*λcI857 ind1 Sam7 nin5 lac*UV5-T7gene1) *dcm* (DE3)	Novagen
Acinetobacter baylyi ADP1	Wild-type; *n*-alkane degrading strain	(9)
ADP1Δ*almA*	Strain ADP1 with *almA* deletion	This study
ADP1Δ*almA* [*almA*$_{tru5aa}$]	Strain ADP1Δ*almA* complemented with *almA*$_{tru5AA}$	This study
Plasmids		
pET-28a(+)	Expression vector; Km^r	Novagen
pET28a-*almA*$_{tru5aa}$	Km^r; pET-28a derivative for *almA*$_{tru5AA}$ expression	This study
Primers[a]		
F(LA_*almA*)	CGTGTTGCTGAGCCTTTAATCA	Forward primer (FP) for left homology arm (LHA) of *almA*
R(LA_*almA*)	cggactggctttctacgtATTGTTTTTATAACCTTGAGAACAATTCT	Reverse primer (RP) for LHA of *almA*
F(Km^r)	atACGTAGAAAGCCAGTCCGCA	FP for kanamycin resistance gene
R(Km^r)	gctcaagTCAGAAGAACTCGTCAAGAAGGC	RP for kanamycin resistance gene
F(RA_*almA*)	cgagttcttctgaCTTGAGCTAAACAAAAAAAGGAGCT	FP for right homology arm (RHA) of *almA*
R(RA_*almA*)	TGGAAGGTATGGGTTTTATGATGC	RP for RHA of *almA*
F(LA_*almA*$_{tru5aa}$)	CGTGTTGCTGAGCCTTTAATCACT	FP for LHA of *almA*$_{tru5aa}$
R(LA_*almA*$_{tru5aa}$)	gcacctacaattaagatatccatATTGTTTTTATAACCTTGAGAACAATTCTATTTTATTAAAACACAAA	RP for LHA of *almA*$_{tru5aa}$
F(*almA*$_{tru5aa}$)	ttctcaaggttataaaaacaatATGGATATCTTAATTGTAGGTGCAGGGATT	FP for *almA*$_{tru5aa}$
R(*almA*$_{tru5aa}$)	ttgaatgtatttagaaaaataaacaaaTTAGGATACTAATTTGGGTTTTGTTTTACGCTCA	RP for *almA*$_{tru5aa}$
F(Amp^r)	aaacaaaacccaaattagtatcctaaTTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCAT	FP for ampicillin resistance gene
R(Amp^r)	ttttttgtttagctcaagTTACCAATGCTTAATCAGTGAGGCACCT	RP for ampicillin resistance gene
F(RA_ *almA*$_{tru5aa}$)	actgattaagcattggtaaCTTGAGCTAAACAAAAAAAGGAGCTTAAACAAAG	FP for RHA of *almA*$_{tru5aa}$
R(RA_ *almA*$_{tru5aa}$)	TGGAAGGTATGGGTTTTATGATGCTAATGAAG	RP for RHA of *almA*$_{tru5aa}$
F(PET_*almA*$_{tru5aa}$)	cgcgcggcagccatatgGATATCTTAATTGTAGGTGCAGGGATTTCGG	FP for *almA*$_{tru5aa}$ expression
R(PET_*almA*$_{tru5aa}$)	gctcgaattcggatccTTAGGATACTAATTTGGGTTTTGTTTTACGCTCAAC	RP for *almA*$_{tru5aa}$ expression
F(D56A)	CCTGGGATTCGTTCTGCGTCTGATATGTCGACT	FP for mutant D56A expression
R(D56A)	AGTCGACATATCAGACGCAGAACGAATCCCAGG	RP for mutant D56A expression
F(R292A)	AACCCATGGGATCAGGCGTTGTGCGTTGTGCCA	FP for mutant R292A expression
R(R292A)	TGGCACAACGCACAACGCCTGATCCCATGGGTT	RP for mutant R292A expression

[a]Lowercase letters in the primers indicate overlapped sequences and underlined regions in the primers indicate mutant sites.

procedure (35). For protein expression and purification, the strain was grown in LB supplemented with 50 µg/mL kanamycin at 37°C and 180 rpm. Upon reaching an OD_{600} of 0.6, the culture was induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside at 16°C overnight (16 h) at 150 rpm. After, the cells were washed and resuspended in phosphate buffer [PB; 100 mM Na_2HPO_4/NaH_2PO_4, pH 7.4; 200 mM NaCl, and 10% (vol/vol) glycerol] and subjected to sonication. The samples were then centrifuged at 13,000 × *g* for 40 min at 4°C and filtered with 0.45 µm filter membranes to obtain the cell extracts. Protein purification was conducted using the ÄKTA start system (GE Healthcare) equipped with a 5-mL HisTrap HP column (GE Healthcare). After sequentially washing the column with 25, 50, and 80 mM imidazole, the recombinant His_6-tagged proteins were eluted with 250 mM imidazole dissolved in PB. Subsequently, imidazole was removed from the protein solution by ultrafiltration. The recombinant His_6-tagged proteins were then separated using sodium dodecyl sulfate-polyacrylamide gel electrophoresis and quantified using a Nano-300 spectrophotometer (Allsheng Instruments Co., Ltd., Hangzhou, China) at A280 nm. The expression plasmids of the mutants $AlmA_{tru5aa}$-D56A and $AlmA_{tru5aa}$-R292A were constructed by site-directed mutagenesis using plasmid pET-*almA*$_{tru5aa}$ as a template and primers listed in Table 1. The mutant proteins were purified using the aforementioned method.

Enzyme activity assay

The enzymatic activity of $AlmA_{tru5aa}$ was assessed using a previously described method (13, 16) with slight modifications. The reaction mixture consisted of 50 mM Tris-HCl (pH 7.4), 1 mM *n*-alkane or aliphatic ketones (as indicated in Table S1), 1 mM NADPH, and 1 mg purified enzyme. The control mixture consisted of the same components without the enzyme. The mixtures were incubated for 10 min at 60°C with shaking and then extracted with hexane and derivatized with trimethylsilyl (if necessary). The reaction products were analyzed via GC–MS.

Gas chromatography–mass spectrometry

The products of enzymatic reactions were analyzed using a GC–MS system (Agilent, USA) consisting of 7890B-GC equipped with an HP-5MS separation column (30 m × 0.25 mm × 0.25 µm) and 5977B-MS detector, as described previously (16). The peaks were identified by comparing the retention times with those of authentic standards and mass spectra to the profiles in the National Institute of Standards and Technology database.

Sequence analysis of AlmA

Functionally identified homologs of AlmA were retrieved from the Uniprot/Swiss-Prot and Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) databases using BLASTP search at default settings. The homologs were manually selected and classified as BVMOs and hydroxylases, according to the references cited. The homologs were aligned using Clustal Omega (28) and processed using ESPript (29). The phylogenetic tree was constructed by the ML method using MEGA v11 software (27) with 1,000 bootstrap replicates and visualized using iTOL (37).

Distribution analysis of canonical *n*-alkane oxygenases in bacterial genomes

All of the available bacterial proteomes (up to May 2023) with whole-genome assemblies (N = 32,796) were downloaded from the National Center for Biotechnology Information RefSeq database for the construction of a microbial proteomic database (N = 122,214,280 protein sequences). This database was then queried against functionally identified protein sequences of AlmA (22), AlkB (18), LadA (21), and CYP153 (17) using BLASTP. Homologous sequences with >40% identity and >70% coverage (38) were retrieved for the distribution and coexistence analysis in bacterial genomes.

Molecular docking simulation

The predicted 3D structure of AlmA with high confidence (pLDDT >85) was obtained from the AlphaFold Protein Structure Database (39) and visualized using the PyMOL Molecular Graphics System v2.5.2 (Schrödinger, LLC) (40). The model with cofactor FAD was predicted by AlphaFill based on sequence and structure similarity (41). Blind docking of the ligands to AlmA model was carried out using CB-Dock2 (42).

ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China (2021YFA0910300). The authors are also grateful to Prof. Xin Song in the Institute of Soil Science, Chinese Academy of Sciences.

AUTHOR AFFILIATIONS

[1]State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic & Developmental Sciences, and School of Life Sciences & Biotechnology, Shanghai Jiao Tong University, Shanghai, China
[2]College of Engineering, Peking University, Beijing, China

AUTHOR ORCIDs

Chao-Fan Yin http://orcid.org/0000-0002-6905-0909
Tao Li http://orcid.org/0000-0002-8255-7798
Ning-Yi Zhou http://orcid.org/0000-0002-0917-5750

FUNDING

Funder	Grant(s)	Author(s)
MOST | National Key Research and Development Program of China (NKPs)	2021YFA0910300	Tao Li

AUTHOR CONTRIBUTIONS

Chao-Fan Yin, Conceptualization, Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft, Writing – review and editing | Yong Nie, Methodology, Validation, Writing – review and editing | Tao Li, Funding acquisition, Resources, Visualization, Writing – review and editing | Ning-Yi Zhou, Conceptualization, Funding acquisition, Project administration, Resources, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

All data used in this study were retrieved from accessible database.

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Table S1; Fig S1-S5 (AEM01625-23-s0001.docx). Supplemental material.

REFERENCES

1. Vasudevan N, Rajaram P. 2001. Bioremediation of oil sludge-contaminated soil. Environ Int 26:409–411. https://doi.org/10.1016/s0160-4120(01)00020-4
2. Zhang J, Gao H, Xue Q. 2020. Potential applications of microbial enhanced oil recovery to heavy oil. Crit Rev Biotechnol 40:459–474. https://doi.org/10.1080/07388551.2020.1739618
3. Saravanan A, Kumar PS, Vardhan KH, Jeevanantham S, Karishma SB, Yaashikaa PR, Vellaichamy P. 2020. A review on systematic approach for microbial enhanced oil recovery technologies: opportunities and challenges. J Clean Prod 258:120777. https://doi.org/10.1016/j.jclepro.2020.120777
4. Zhang C, Wu D, Ren H. 2020. Bioremediation of oil contaminated soil using agricultural wastes via microbial consortium. Sci Rep 10:9188. https://doi.org/10.1038/s41598-020-66169-5
5. Wentzel A, Ellingsen TE, Kotlar H-K, Zotchev SB, Throne-Holst M. 2007. Bacterial metabolism of long-chain *n*-alkanes. Appl Microbiol Biotechnol 76:1209–1221. https://doi.org/10.1007/s00253-007-1119-1
6. Van Beilen JB, Li Z, Duetz WA, Smits THM, Witholt B. 2003. Diversity of alkane hydroxylase systems in the environment. Oil Gas Sci Technol 58:427–440. https://doi.org/10.2516/ogst:2003026
7. Ratajczak A, Geissdörfer W, Hillen W. 1998. Alkane hydroxylase from *Acinetobacter* sp. strain ADP1 is encoded by *alkM* and belongs to a new family of bacterial integral-membrane hydrocarbon hydroxylases. Appl Environ Microbiol 64:1175–1179. https://doi.org/10.1128/AEM.64.4.1175-1179.1998
8. Metzgar D, Bacher JM, Pezo V, Reader J, Döring V, Schimmel P, Marlière P, de Crécy-Lagard V. 2004. *Acinetobacter* sp. ADP1: an ideal model organism for genetic analysis and genome engineering. Nucleic Acids Res 32:5780–5790. https://doi.org/10.1093/nar/gkh881
9. Barbe V, Vallenet D, Fonknechten N, Kreimeyer A, Oztas S, Labarre L, Cruveiller S, Robert C, Duprat S, Wincker P, Ornston LN, Weissenbach J, Marlière P, Cohen GN, Médigue C. 2004. Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile and naturally transformation competent bacterium. Nucleic Acids Res 32:5766–5779. https://doi.org/10.1093/nar/gkh910
10. Rojo F. 2009. Degradation of alkanes by bacteria. Environ Microbiol 11:2477–2490. https://doi.org/10.1111/j.1462-2920.2009.01948.x
11. Li L, Liu X, Yang W, Xu F, Wang W, Feng L, Bartlam M, Wang L, Rao Z. 2008. Crystal structure of long-chain alkane monooxygenase (LadA) in complex with coenzyme FMN: unveiling the long-chain alkane hydroxylase. J Mol Biol 376:453–465. https://doi.org/10.1016/j.jmb.2007.11.069
12. van Beilen JB, Funhoff EG. 2007. Alkane hydroxylases involved in microbial alkane degradation. Appl Microbiol Biotechnol 74:13–21. https://doi.org/10.1007/s00253-006-0748-0
13. Wang W, Shao Z. 2014. The long-chain alkane metabolism network of *Alcanivorax dieselolei*. Nat Commun 5:5755. https://doi.org/10.1038/ncomms6755
14. Nie Y, Chi C-Q, Fang H, Liang J-L, Lu S-L, Lai G-L, Tang Y-Q, Wu X-L. 2014. Diverse alkane hydroxylase genes in microorganisms and environments. Sci Rep 4:4968. https://doi.org/10.1038/srep04968
15. Wang W, Shao Z. 2013. Enzymes and genes involved in aerobic alkane degradation. Front Microbiol 4:116. https://doi.org/10.3389/fmicb.2013.00116
16. Yin CF, Xu Y, Li T, Zhou NY. 2022. Wide distribution of the *sad* gene cluster for sub - terminal oxidation in alkane utilizers. Environ Microbiol 24:6307–6319. https://doi.org/10.1111/1462-2920.16124
17. Maier T, Förster HH, Asperger O, Hahn U. 2001. Molecular characterization of the 56-kDa CYP153 from *Acinetobacter* sp. EB104. Biochem

Biophys Res Commun 286:652–658. https://doi.org/10.1006/bbrc.2001.5449
18. van Beilen JB, Wubbolts MG, Witholt B. 1994. Genetics of alkane oxidation by *Pseudomonas oleovorans*. Biodegradation 5:161–174. https://doi.org/10.1007/BF00696457
19. Guo X, Zhang J, Han L, Lee J, Williams SC, Forsberg A, Xu Y, Austin RN, Feng L. 2023. Structure and mechanism of the alkane-oxidizing enzyme AlkB. Nat Commun 14:2180. https://doi.org/10.1038/s41467-023-37869-z
20. Tani A, Ishige T, Sakai Y, Kato N. 2001. Gene structures and regulation of the alkane hydroxylase complex in *Acinetobacter* sp. strain M-1. J Bacteriol 183:1819–1823. https://doi.org/10.1128/JB.183.5.1819-1823.2001
21. Feng L, Wang W, Cheng J, Ren Y, Zhao G, Gao C, Tang Y, Liu X, Han W, Peng X, Liu R, Wang L. 2007. Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. Proc Natl Acad Sci U S A 104:5602–5607. https://doi.org/10.1073/pnas.0609650104
22. Throne-Holst M, Wentzel A, Ellingsen TE, Kotlar H-K, Zotchev SB. 2007. Identification of novel genes involved in long-chain *n*-alkane degradation by *Acinetobacter* sp. strain DSM 17874. Appl Environ Microbiol 73:3327–3332. https://doi.org/10.1128/AEM.00064-07
23. Minerdi D, Zgrablic I, Sadeghi SJ, Gilardi G. 2012. Identification of a novel Baeyer-Villiger monooxygenase from *Acinetobacter radioresistens*: close relationship to the *Mycobacterium tuberculosis* prodrug activator EtaA. Microb Biotechnol 5:700–716. https://doi.org/10.1111/j.1751-7915.2012.00356.x
24. Leisch H, Morley K, Lau PCK. 2011. Baeyer-Villiger monooxygenases: more than just green chemistry. Chem Rev 111:4165–4222. https://doi.org/10.1021/cr1003437
25. Fraaije MW, Kamerbeek NM, van Berkel WJH, Janssen DB. 2002. Identification of a Baeyer-Villiger monooxygenase sequence motif. FEBS Lett 518:43–47. https://doi.org/10.1016/s0014-5793(02)02623-6
26. Yachnin BJ, McEvoy MB, MacCuish RJD, Morley KL, Lau PCK, Berghuis AM. 2014. Lactone-bound structures of cyclohexanone monooxygenase provide insight into the stereochemistry of catalysis. ACS Chem Biol 9:2843–2851. https://doi.org/10.1021/cb500442e
27. Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol 38:3022–3027. https://doi.org/10.1093/molbev/msab120
28. Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A, Lopez R. 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res 50:W276–W279. https://doi.org/10.1093/nar/gkac240
29. Gouet P, Courcelle E, Stuart DI, Métoz F. 1999. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15:305–308. https://doi.org/10.1093/bioinformatics/15.4.305
30. Fürst M, Gran-Scheuch A, Aalbers FS, Fraaije MW. 2019. Baeyer–Villiger monooxygenases: tunable oxidative biocatalysts. ACS Catal 9:11207–11241. https://doi.org/10.1021/acscatal.9b03396
31. Inderthal H, Tai SL, Harrison STL. 2021. Non-hydrolyzable plastics - an Interdisciplinary look at plastic bio-oxidation. Trends Biotechnol 39:12–23. https://doi.org/10.1016/j.tibtech.2020.05.004
32. Liu C, Wang W, Wu Y, Zhou Z, Lai Q, Shao Z. 2011. Multiple alkane hydroxylase systems in a marine alkane degrader, *Alcanivorax dieselolei* B-5. Environ Microbiol 13:1168–1178. https://doi.org/10.1111/j.1462-2920.2010.02416.x
33. Viggor S, Jõesaar M, Vedler E, Kiiker R, Pärnpuu L, Heinaru A. 2015. Occurrence of diverse alkane hydroxylase *alkB* genes in indigenous oil-degrading bacteria of baltic sea surface water. Mar Pollut Bull 101:507–516. https://doi.org/10.1016/j.marpolbul.2015.10.064
34. Liu H, Wang SJ, Zhou NY. 2005. A new isolate of *Pseudomonas stutzerithat* degrades 2-chloronitrobenzene. Biotechnol Lett 27:275–278. https://doi.org/10.1007/s10529-004-8293-3
35. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold spring harbor laboratory press.
36. Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, Pühler A. 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145:69–73. https://doi.org/10.1016/0378-1119(94)90324-7
37. Letunic I, Bork P. 2019. Interactive tree of life (iTOL) V4: recent updates and new developments. Nucleic Acids Res 47:W256–W259. https://doi.org/10.1093/nar/gkz239
38. Addou S, Rentzsch R, Lee D, Orengo CA. 2009. Domain-based and family-specific sequence identity thresholds increase the levels of reliable protein function transfer. J Mol Biol 387:416–430. https://doi.org/10.1016/j.jmb.2008.12.045
39. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, et al. 2022. Alphafold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res 50:D439–D444. https://doi.org/10.1093/nar/gkab1061
40. Schrodinger LLC. 2015. The PyMOL molecular graphics system version, 1.8
41. Hekkelman ML, de Vries I, Joosten RP, Perrakis A. 2023. AlphaFill: enriching AlphaFold models with ligands and cofactors. Nat Methods 20:205–213. https://doi.org/10.1038/s41592-022-01685-y
42. Liu Y, Yang X, Gan J, Chen S, Xiao ZX, Cao Y. 2022. CB-Dock2: improved protein-ligand blind docking by integrating cavity detection, docking and homologous template fitting. Nucleic Acids Res 50:W159–W164. https://doi.org/10.1093/nar/gkac394