



The plant pathogen enzyme AldC is a long-chain aliphatic aldehyde dehydrogenase

Received for publication, June 15, 2020, and in revised form, August 11, 2020. Published, Papers in Press, August 12, 2020. DOI 10.1074/jbc.RA120.014747

Soon Goo Lee^{1,2,†}, Kate Harline^{1,†}, Orchid Abar¹, Sakirat O. Akadri¹, Alexander G. Bastian¹, Hui-Yuan S. Chen¹, Michael Duan¹, Caroline M. Focht¹, Amanda R. Groziak¹, Jesse Kao¹, Jagdeesh S. Kottapalli¹, Matthew C. Leong¹, Joy J. Lin¹, Regina Liu¹, Joanna E. Luo¹, Christine M. Meyer¹, Albert F. Mo¹, Seong Ho Pahng¹, Vinay Penna¹, Chris D. Raciti¹, Abhinav Srinath¹, Shwetha Sudhakar¹, Joseph D. Tang¹, Brian R. Cox², Cynthia K. Holland^{1,3}, Barrie Cascella¹, Wilhelm Cruz¹, Sheri A. McClerkin^{1,4}, Barbara N. Kunkel¹, and Joseph M. Jez^{1,*}

From the ¹Department of Biology, Washington University in St. Louis, St. Louis, Missouri, USA, the ²Department of Chemistry and Biochemistry, University of North Carolina–Wilmington, Wilmington, North Carolina, USA, the ³Department of Biology, Williams College, Williamstown, Massachusetts, USA, and the ⁴Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, Illinois, USA

Edited by Chris Whitfield

Aldehyde dehydrogenases are versatile enzymes that serve a range of biochemical functions. Although traditionally considered metabolic housekeeping enzymes because of their ability to detoxify reactive aldehydes, like those generated from lipid peroxidation damage, the contributions of these enzymes to other biological processes are widespread. For example, the plant pathogen *Pseudomonas syringae* strain PtoDC3000 uses an indole-3-acetaldehyde dehydrogenase to synthesize the phytohormone indole-3-acetic acid to elude host responses. Here we investigate the biochemical function of AldC from PtoDC3000. Analysis of the substrate profile of AldC suggests that this enzyme functions as a long-chain aliphatic aldehyde dehydrogenase. The 2.5 Å resolution X-ray crystal of the AldC C291A mutant in a dead-end complex with octanal and NAD⁺ reveals an apolar binding site primed for aliphatic aldehyde substrate recognition. Functional characterization of site-directed mutants targeting the substrate- and NAD(H)-binding sites identifies key residues in the active site for ligand interactions, including those in the “aromatic box” that define the aldehyde-binding site. Overall, this study provides molecular insight for understanding the evolution of the prokaryotic aldehyde dehydrogenase superfamily and their diversity of function.

In all organisms, the diversification of large superfamilies of enzymes provides a foundation for the evolution of biochemical capacity and the ability to metabolize varied small molecules (1). Typically, across a superfamily, the core chemical mechanism is retained with substrate profiles becoming either specialized or promiscuous, depending on evolution and the metabolic needs of the organism (2). Classic examples of enzyme superfamilies found across all kingdoms include the cytochromes P450, alcohol dehydrogenases, aldo-keto reductases, and aldehyde dehydrogenases (3–10). For example, aldehyde dehydrogenases are NAD(P)(H)-dependent enzymes that metabolize a wide range of aldehydes to their corresponding carboxylic acids in prokaryotes and eukaryotes and tend to

be encoded by multiple genes in the genome of a given species (9, 10) (Fig. 1).

Aldehyde dehydrogenases are generally associated with the detoxification of aldehydes, which are highly reactive compounds generated through cellular metabolism (11). For example, these enzymes can scavenge aldehydes, such as malondialdehyde resulting from lipid peroxidation, and convert them to a less chemically reactive carboxylic acid (12). Aldehyde dehydrogenases and their biochemical functions are also linked to a wide variety of biochemical processes ranging from ethanol metabolism via oxidation of acetaldehyde into acetate (13–15), polyamine metabolism (16), and plant cell wall ester biogenesis (17, 18) to protective cellular responses to stresses such as dehydration, osmotic shock, and temperature changes (19, 20). Recent work has also linked aldehyde dehydrogenase activity to the synthesis of indole-3-acetic acid, the primary plant hormone auxin, in the plant pathogenic microbe *Pseudomonas syringae* strain PtoDC3000 (21).

To suppress host defenses and promote diseases development, *Pseudomonas syringae* produces a variety of virulence factors, including phytohormones or chemical mimics of hormones, to manipulate hormone signaling in its host plants (22–24). *P. syringae* and many other plant-associated microbial pathogens can synthesize the major auxin indole-3-acetic acid (IAA), whose production is implicated in pathogen virulence (21, 25–29). PtoDC3000 was shown to synthesize IAA using an uncharacterized pathway requiring indole-3-acetaldehyde dehydrogenase activity (21). Previously, a mutation in *Azospirillum brasilense* (*aldA*) was identified that decreased IAA production and was linked to a gene and annotated as encoding an aldehyde dehydrogenase (30). Bioinformatic analysis of PtoDC3000 identified a set of aldehyde dehydrogenases (AldA–C) sharing 30–40% amino acid identity with each other. Subsequent metabolic, biochemical, and *in planta* analyses of AldA (UniProt: PSPTO_0092), AldB (UniProt: PSPTO_2673), and AldC (UniProt: PSPTO_3644) demonstrated that AldA functions as an indole-3-acetaldehyde dehydrogenase, is essential for IAA synthesis, and contributes to virulence of PtoDC3000 (21). AldB may also contribute to

This article contains supporting information.

[†]These authors contributed equally to this work.

*For correspondence: Joseph M. Jez, jjez@wustl.edu.

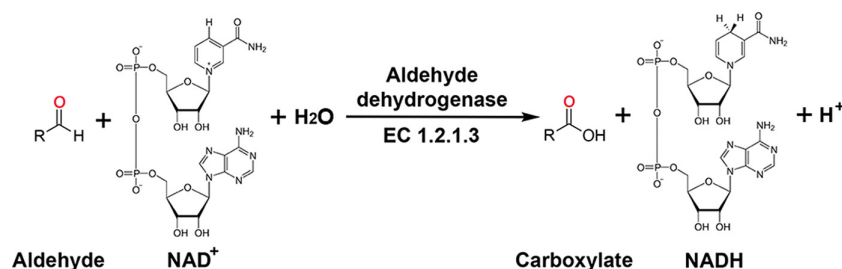


Figure 1. Overall reaction catalyzed by NAD(P)(H)-dependent aldehyde dehydrogenases.

IAA synthesis but is not as metabolically or kinetically efficient as AldA for IAA production (21). Analysis of AldC indicates that it lacks a role in IAA synthesis (21); however, its potential biochemical role in *Pto*DC3000 was not fully examined.

Here we investigate the substrate profile of AldC from *Pto*DC3000, which suggests that this enzyme functions as a long-chain aliphatic aldehyde dehydrogenase. The 2.5 Å resolution X-ray crystal of the AldC C291A mutant in complex with octanal and NAD⁺ and biochemical analysis of a set of site-directed mutants provide insight on substrate recognition in this enzyme. Comparison of the three-dimensional structures of AldA and AldC from *Pto*DC3000 reveals the sequence and structural changes that lead to distinct substrate profiles of these aldehyde dehydrogenases.

Results and discussion

Biochemical analysis of AldC as a long-chain aliphatic aldehyde dehydrogenase

Previously, three aldehyde dehydrogenases (AldA–C) from the plant pathogen *P. syringae* strain *Pto*DC3000 were identified and examined for their contribution to the synthesis of IAA and virulence (21). Each protein was expressed and purified to homogeneity for enzyme assays using indole-3-acetaldehyde as a substrate (21). Unlike AldA and AldB, which functioned as physiological tetramers, AldC was shown to be a homodimer (21). Steady-state kinetic parameters for AldA using indole-3-acetaldehyde showed this enzyme to be 130- and 710-fold more efficient than AldB and AldC, respectively (21).

Bioinformatic analysis predicted that AldC contains the PF00171 domain, a signature domain of the aldehyde dehydrogenase superfamily, ranging from Ser³³ to Ile⁴⁸² (Fig. 2A). Specifically, the 4 catalytic residues (Asn¹⁵⁹, Glu²⁵⁷, Gly²⁸⁸, and Cys²⁹¹) and 19 residues in the NAD⁺-binding site (Ile¹⁵³–Asn¹⁵⁹, Lys¹⁸², Gly²¹⁹, Ile²³³–Ser²³⁶, Ala²³⁹, Leu²⁴², Glu²⁵⁷, Leu²⁵⁸, Gly²⁵⁹, Cys²⁹¹, Glu³⁹¹, and Phe³⁹³) define the aldehyde dehydrogenase consensus sequence motifs of AldC. The NCBI Conserved Protein Domain Family server places AldC in the cd01738/ALDH_CddD_SSP0762 family, 1 of 42 aldehyde dehydrogenase families found in the *Pseudomonas* (31, 32).

The *Pseudomonas* orthologous groups classification system in the *Pseudomonas* Genome Database (RRID:SCR_006590) found orthologs of AldC (group ID POG018413) from 93 *Pseudomonas* species and strains, including *P. aeruginosa*, *P. putida*,

P. fluorescens, and *P. savastanoi* (Fig. 2B and Table S1). AldC belongs to a clade consisting of aldehyde dehydrogenases, which share ~90% sequence identity, from various plant pathogenic *Pseudomonas* spp.: *P. syringae*, *P. viridiflava*, *P. savastanoi*, and *P. amygdali* in the *P. syringae* phylogenetic group. None of AldC-related enzymes from *Pseudomonas* had previously been experimentally characterized, and their substrates and functions had yet to be described.

To determine the substrate preference of AldC, a panel of 23 molecules (Table S2), including short- to long-chain aliphatic aldehydes (acetaldehyde, propionaldehyde, butyraldehyde, valeraldehyde (pentanal), isovaleraldehyde, 3-methylcrotonaldehyde, hexanal, *trans*-2-hexen-1-al, heptanal, octanal, nonanal, *trans*-2-nonenal), betaine aldehyde, and aromatic aldehydes (benzaldehyde, *m*-anisaldehyde, *p*-anisaldehyde, phenylacetaldehyde, cinnamaldehyde, hydrocinnamaldehyde, coniferylaldehyde, sinapaldehyde, 4-pyridinecarboxaldehyde, and indole-3-acetaldehyde) was used to screen for enzymatic activity. Spectrophotometric assays of AldC identified aliphatic aldehydes of 5–9-carbon length, as well as hydrocinnamaldehyde and 4-pyridinecarboxaldehyde, as substrates (Fig. 3A) with octanal having the highest specific activity (Fig. 3B). These data suggest that short 2–4-carbon aldehydes, branched aliphatic aldehydes, and larger aromatic aldehydes are poor substrates for AldC. To evaluate the nicotinamide cofactor preference of AldC, the activity of the enzyme was tested with octanal and either NAD⁺ or NADP⁺, which showed a distinct preference for NAD⁺ (Fig. 3C).

Steady-state kinetic analysis of AldC with valeraldehyde, hexanal, heptanal, octanal, and nonanal indicates that the 8-carbon substrate (*i.e.* octanal) is the preferred aliphatic aldehyde substrate (Table 1). Although the k_{cat} values of AldC vary less than 3-fold between these substrates, the K_m value for octanal is the lowest (1.2 mM) with that of valeraldehyde as the highest (48.8 mM). The overall effect is a marked difference in catalytic efficiency (k_{cat}/K_m) between octanal and the other aliphatic substrates ranging from use of nonanal exhibiting a 10-fold reduction in k_{cat}/K_m to ~80-fold less efficient use of the 5-carbon substrate valeraldehyde. Generally, the aromatic aldehyde substrates hydrocinnamaldehyde and indole-3-acetaldehyde were poorly used by AldC with catalytic efficiencies comparable with valeraldehyde and 4-pyridinecarboxaldehyde used with a k_{cat}/K_m value 6-fold lower than octanal (Table 1). Overall, the biochemical analysis of AldC suggests that this enzyme functions primarily as a long-chain aliphatic

Analysis of AldC from *P. syringae* DC3000

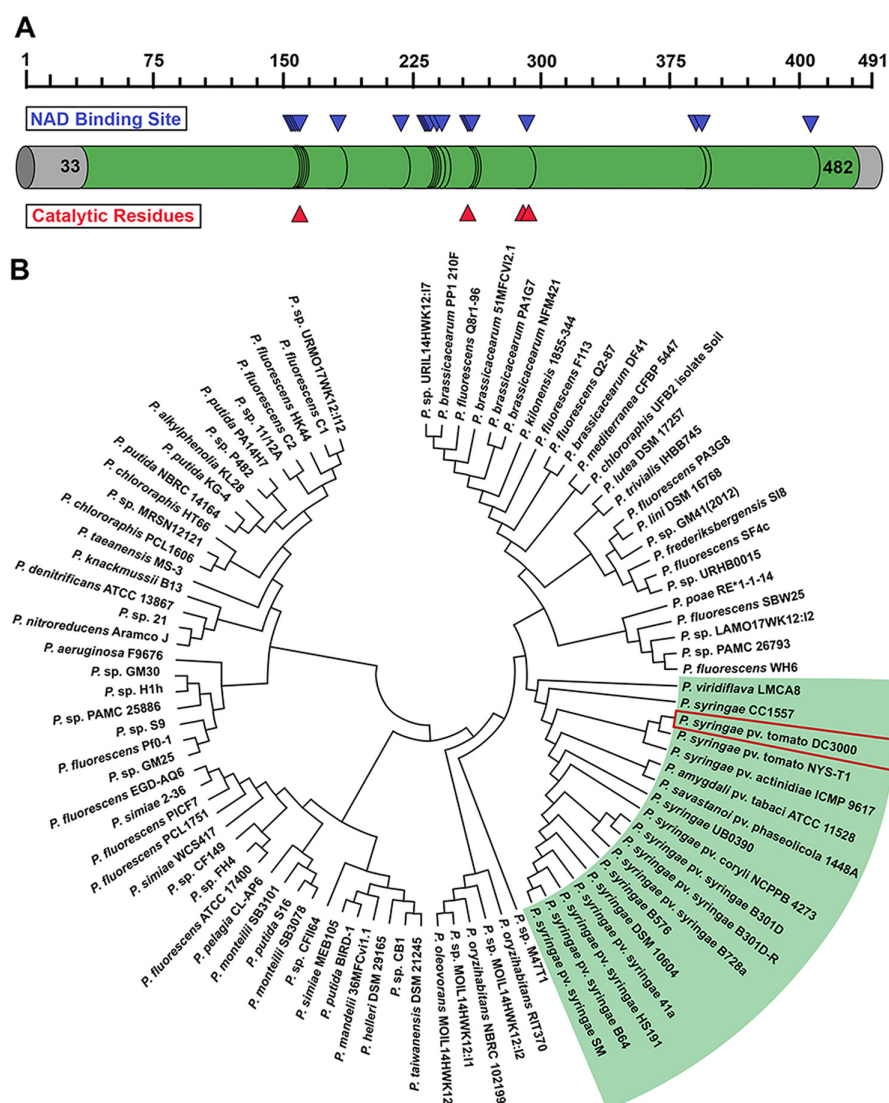


Figure 2. Domain architecture and phylogeny of AldC from *P. syringae* strain PtoDC3000. A, the UniProt aldehyde dehydrogenase (ALDH) domain (green), NAD(P)(H)-binding site (blue triangles), and catalytic residues (red triangles) of PtoDC3000 AldC (WP_011104646.1) were predicted using InterPro79.0 (59). B, the PtoDC3000 AldC sequence was used as a BLAST query to identify ortholog group members of *aldC* (group ID POG018413) from the *Pseudomonas* Orthologous Groups classification system in the *Pseudomonas* Genome Database (RRID:SCR_006590) (Table S1). Amino acid sequences from 93 *Pseudomonas* species strains were multialigned using Clustal Omega (60). The phylogenetic tree was generated with MEGAX (61) with evolutionary relationships inferred using the maximum likelihood method and JTT matrix-based model (62). A clade consisting of aldehyde dehydrogenases from plant pathogenic *Pseudomonas* species and strains is highlighted (green) with AldC from PtoDC3000 indicated.

aldehyde dehydrogenase (Fig. 3D). Moreover, conservation of related homologs in 93 different *Pseudomonas* species and strains (Fig. 2B) indicates a likely conserved function across these organisms.

Three-dimensional structure of AldC C291A mutant in complex with octanal and NAD⁺

To understand how AldC recognizes aliphatic aldehyde substrates, crystals of the AldC C291A mutant were grown in the presence of octanal and NAD⁺. The analogous cysteine to Cys²⁹¹ of AldC in aldehyde dehydrogenases, including AldA from *P. syringae*, is the essential catalytic residue (9, 10, 14–16,

21). The AldC C291A point mutant was generated by PCR mutagenesis, and the resulting protein was expressed, purified, and crystallized with the goal of obtaining a dead-end complex of the enzyme with octanal and NAD⁺ bound in the active site.

The 2.5 Å resolution X-ray crystal structure of the AldC (C291A)·octanal·NAD⁺ complex was solved by molecular replacement using the three-dimensional structure of *P. syringae* AldA as a search model (21) (Table 2). Each of the two unique protein chains in the asymmetric unit form a corresponding physiological dimer through crystallographic symmetry (Fig. 4A). The secondary structure features and domains of the AldC monomer are similar to those of other aldehyde dehydrogenase family members (Fig. 4, B and C). The N-terminal

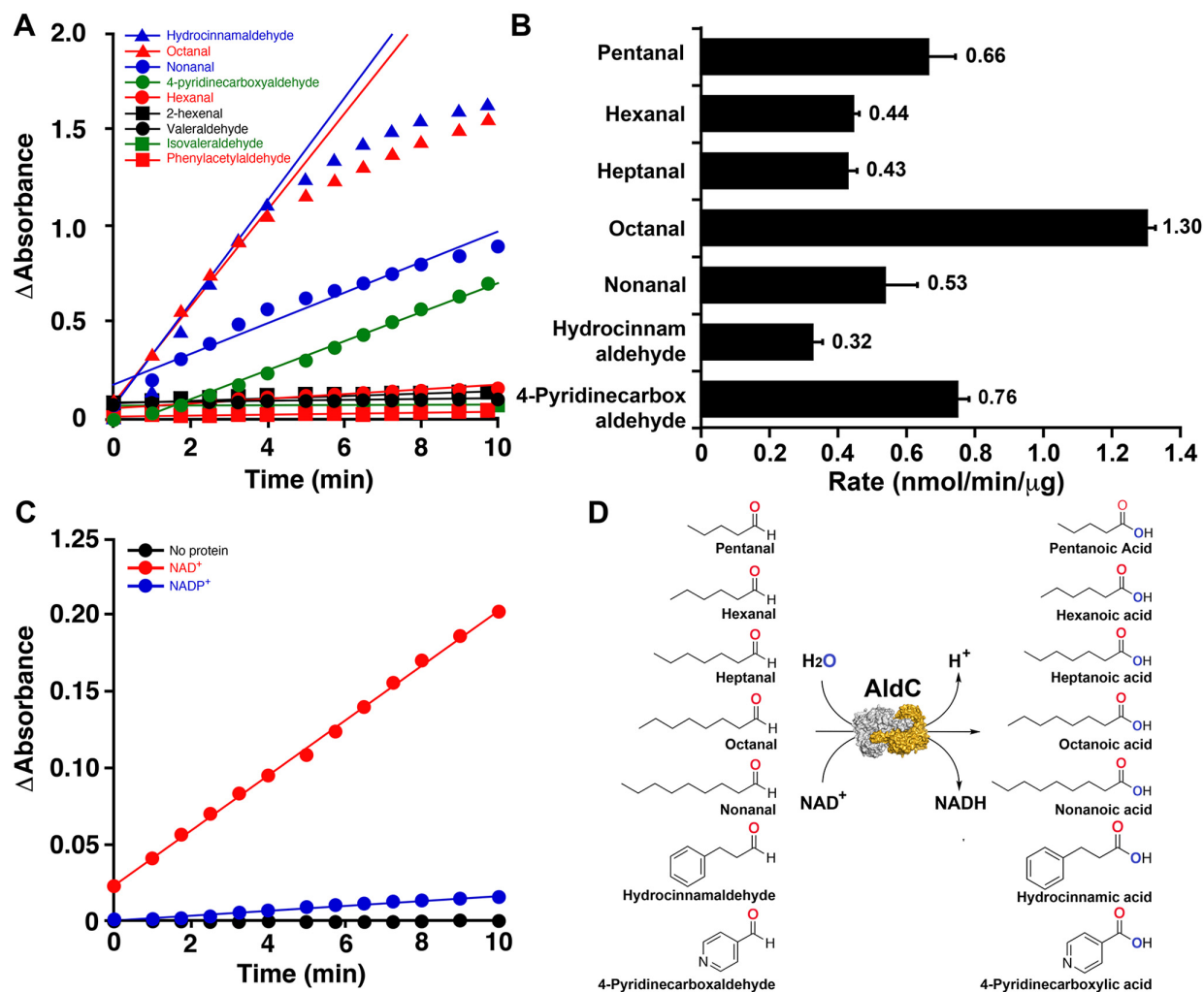


Figure 3. Screening of AldC with aldehyde substrates. A, screening of aldehyde substrates. Assays were performed as described under “Experimental procedures.” Enzymatic activity was measured spectrophotometrically ($A_{340\text{ nm}}$) in assays with a fixed concentration of NAD⁺ (2 mM) and each aldehyde (5 mM). Spectrophotometric absorbance changes *versus* time are shown for AldC-catalyzed conversion of hydrocinnamaldehyde, octanal, nonanal, 4-pyridinecarboxyaldehyde, hexanal, 2-hexenal, valeraldehyde, isovaleraldehyde, and phenylacetylaldehyde. B, comparison of AldC specific activity for substrate aldehydes. Average values $\pm$ S.D. ($n = 3$) are shown. C, AldC nicotinamide cofactor preference. Spectrophotometric data for assays of AldC using octanal (5 mM) and either NAD⁺ and NADP⁺ (each at 2 mM) is shown. D, summary of aldehyde substrates and potential reactions catalyzed by AldC.

Rossmann-fold domain contains a central β -sheet ($\beta 9$ – $\beta 8$ – $\beta 7$ – $\beta 10$ – $\beta 11$) surrounded by α -helices to form the NAD(H)-binding site (Fig. 4, B and C, blue). The C-terminal region consists of a mixed α/β domain, which includes the catalytic cysteine residue and forms the aldehyde-binding site (Fig. 4, B and C, red). An interdomain linker region (Fig. 4, B and C, green) connects the N- and C-terminal domains of AldC. A small three-stranded β -sheet domain (Fig. 4, B and C, gold) facilitates oligomerization.

The overall fold of AldC shares structural similarity with multiple aldehyde dehydrogenases, including *P. syringae* AldA, from a variety of microbes and eukaryotes (Fig. S1 and Table S3). Overlays of the AldC structure with twelve aldehyde, retinal, betaine aldehyde, and indole-3-acetaldehyde (*i.e.* AldA) dehydrogenases (14, 15, 21, 33–37), which are related by 38–46% amino acid identity, show the conservation of the three-dimensional fold with 0.7–1.2 Å RMSD for 400–450 C α -atoms in

these enzymes from widely varied organisms (Fig. S1). Amino acid sequence comparisons of these proteins highlight how the catalytic cysteine and residues of the NAD(H)-binding site are highly conserved with major variations in the substrate-binding site leading to functional differences in these proteins (Fig. S2).

Although the specific determinants of oligomerization in aldehyde dehydrogenases can be highly variable (38), the X-ray crystal structure of AldC provides possible insight on its dimer structure compared with the tetramer reported for AldA (21). The *P. syringae* AldA tetramer is formed from a dimer of dimers with an extensive interaction surface, shown in Fig. 5 (A and B) with the oligomerization region highlighted in the close-ups. Fig. 5 (C and D) shows the same region of the AldC dimer for comparison. Multiple amino acid differences between AldA and AldC in the oligomerization domain alter both the electrostatic surface charge, which is strongly basic in AldA, and the surface topology, especially in the positioning of $\beta 5$, $\beta 6$, and the

Analysis of AldC from *P. syringae* DC3000

Table 1

Steady-state kinetic parameters of AldC with aldehyde substrates

Assays were performed as described under "Experimental procedures." The average values $\pm$ S.D. ($n = 3$) are shown. Assays with varied NAD⁺ used octanal as the substrate.

Substrate	K_{cat}	K_m	K_{cat}/K_m
	min^{-1}	mM	$\text{M}^{-1} \text{s}^{-1}$
Valeraldehyde/pentanal	34.3 ± 4.2	48.8 ± 17.1	12
Hexanal	22.9 ± 1.0	16.4 ± 2.3	23
Heptanal	22.1 ± 1.5	7.4 ± 2.0	49
Octanal	67.4 ± 1.4	1.2 ± 0.1	924
Nonanal	27.7 ± 5.0	4.9 ± 2.3	94
Hydrocinnamaldehyde	16.8 ± 1.6	14.2 ± 3.1	20
4-Pyridinecarboxaldehyde	38.7 ± 1.9	4.4 ± 0.6	147
Indole-3-acetaldehyde ^a	3.6 ± 0.2	1.3 ± 0.3	14
NAD ⁺ (octanal)	30.9 ± 0.4	0.5 ± 0.1	1,050

^aSteady-state kinetic parameters of AldC with indole-3-acetaldehyde were previously reported (21).

Table 2

Summary of crystallographic statistics for AldC structures

Crystal	AldC (C291A)·octanal·NAD ⁺
Space group	C222 ₁
Cell dimensions (Å)	
<i>a</i>	92.19
<i>b</i>	92.17
<i>c</i>	231.4
Data collection	
Wavelength	0.979 Å
Resolution range (highest shell)	39.6–2.52 Å (2.54–2.52 Å)
Reflections (total/unique)	116,137/32,956
Completeness (highest shell)	97.8% (99.1%)
$\langle I/\sigma \rangle$ (highest shell)	12.8 (1.8)
R_{sym}^a (highest shell)	6.0% (52.9%)
Refinement	
$R_{\text{cryst}}^b/R_{\text{free}}^c$	17.1%/21.9%
No. of protein atoms	7,180
No. of waters	197
No. of ligand atoms	106
RMSE	
Bond lengths (Å)	0.008
Bond angles (°)	0.93
Average B-factor (Å ²)	
Protein	45.9
Water	43.7
Ligand	55.9
Stereochemistry (%)	
Favored	95.7
Allowed	4.3
Outliers	0

^a $R_{\text{sym}} = \sum |I_h - \langle I_h \rangle| / \sum I_h$, where $\langle I_h \rangle$ is the average intensity over symmetry.

^b $R_{\text{cryst}} = \sum |F_o - \langle F_o \rangle| / \sum F_o$, where summation is over the data used for refinement.

^c R_{free} is defined the same as R_{cryst} but was calculated using 5% of data excluded from refinement.

$\beta 5$ – $\beta 6$ loop of the oligomerization domain of the two proteins. In addition, the length of the $\beta 5$ – $\beta 6$ loop differs between AldA (Ala¹⁴³–Val¹⁵³) and AldC (Val¹³⁸–Thr¹⁴²), which contributes to alterations in the interaction region (Fig. 5, *B* and *D*, and Fig. S2, *orange box*). In AldC, these changes appear to alter what would be the corresponding tetramer (*i.e.* dimer–dimer) interface of AldA but may only be one contributing factor to different oligomerization.

Active site structure of AldC

Unambiguous electron densities for NAD⁺ and octanal define how these ligands bind to the AldC C291A mutant and indicate the location of the substrate and cofactor-binding sites (Fig. 6A). Forming a dead-end complex, NAD⁺ and octanal

occupy two separate pockets, which are at the opposite ends of a ~ 45 -Å-long tunnel (Fig. 6B). The Rossmann fold of the NAD (H)-binding domain provides extensive polar and apolar interactions that position the nicotinamide ring of NAD⁺ in proximity to the C291A point mutation, which would be the invariant catalytic cysteine in WT AldC (Fig. 6C). The nicotinamide ring is mainly held in place by van der Waals contacts with Leu²⁵⁸, Leu⁴¹⁹, and Phe⁴⁵⁶ and a critical hydrogen bond from the backbone carbonyl of Leu²⁵⁸ to the NH₂ group of the cofactor carboxamide. The nicotinamide-ribose of NAD⁺ establishes multiple polar interactions with the backbone carbonyl of Gly²³⁵ and the carboxylate side chain of Glu³⁹¹. In addition, the adenine ring of NAD⁺ is mainly stabilized by multiple van der Waals interactions with Pro²¹⁶, Ile²³³, Leu²⁴², and Val²⁴³, along with a hydrogen bond with a water molecule in an apolar pocket. As with nicotinamide-ribose binding, polar interactions between the adenine-ribose ring and the side-chains of Lys¹⁸² and Glu¹⁸⁵ contribute to NAD⁺ binding. Interaction of Glu¹⁸⁵ with the 2'-hydroxyl-group of the adenine-ribose is the key structural determinant of the cofactor specificity, as observed in other aldehyde dehydrogenases (39). AldC is not able to accommodate the 2'-phosphate of NADP(H) sterically and electrostatically, which is consistent with its preference for NAD(H) as cofactor (39). Sequence comparison of AldC and structurally related aldehyde dehydrogenases (Fig. S2) underscores the conservation of the NAD(H)-binding site across these enzymes from distantly related organisms.

In contrast to the NAD(H)-binding site, apolar interactions dominate the octanal binding in the hydrophobic substrate-binding pocket (Fig. 6, *B* and *C*). A cluster of aromatic residues (Trp¹⁶⁰, Tyr¹⁶³, Trp⁴⁵⁰, Phe⁴⁵⁶, and Tyr⁴⁶⁸) and two apolar residues (Met¹¹⁴ and Leu¹¹⁸) provide the hydrophobic environment that accommodates octanal and other aliphatic aldehydes. As described for other aldehyde dehydrogenases (40–42), the substrate-binding site forms an "aromatic box" for adaptable apolar ligand interaction.

In the AldC(C291A)·octanal·NAD⁺ complex, the 8-carbon chain of octanal extends toward the solvent exposed, but highly apolar, opening of the aldehyde-binding site. This places the reactive aldehyde group, which hydrogen bonds with Asn¹⁵⁹ and the backbone nitrogen of Ala²⁹¹, in proximity to the C4 position of the nicotinamide ring that undergoes hydride transfer in the chemical reaction and the location of the catalytic cysteine. Multiple sequence alignment of AldC and other aldehyde dehydrogenases (Fig. S2) shows that the residues forming the aldehyde-binding site are highly variable, which account for substrate preferences in these enzymes.

Molecular docking of the valeraldehyde (pentanal), hexanal, heptanal, nonanal, hydrocinnamaldehyde, and 4-pyridinecarboxaldehyde (Fig. S3) provides insight on likely binding modes for each substrate that place the reactive aldehyde near the catalytic center of AldC and the aliphatic or aromatic portions of each into the apolar tunnel leading toward the solvent opening. Although the size of the site would allow short-chain aliphatic substrates, such as acetaldehyde, propionaldehyde, and butyraldehyde, to bind, the region of the substrate-binding site where these molecules would fit is more polar than the largely hydrophobic region of the tunnel accessible to longer-chained

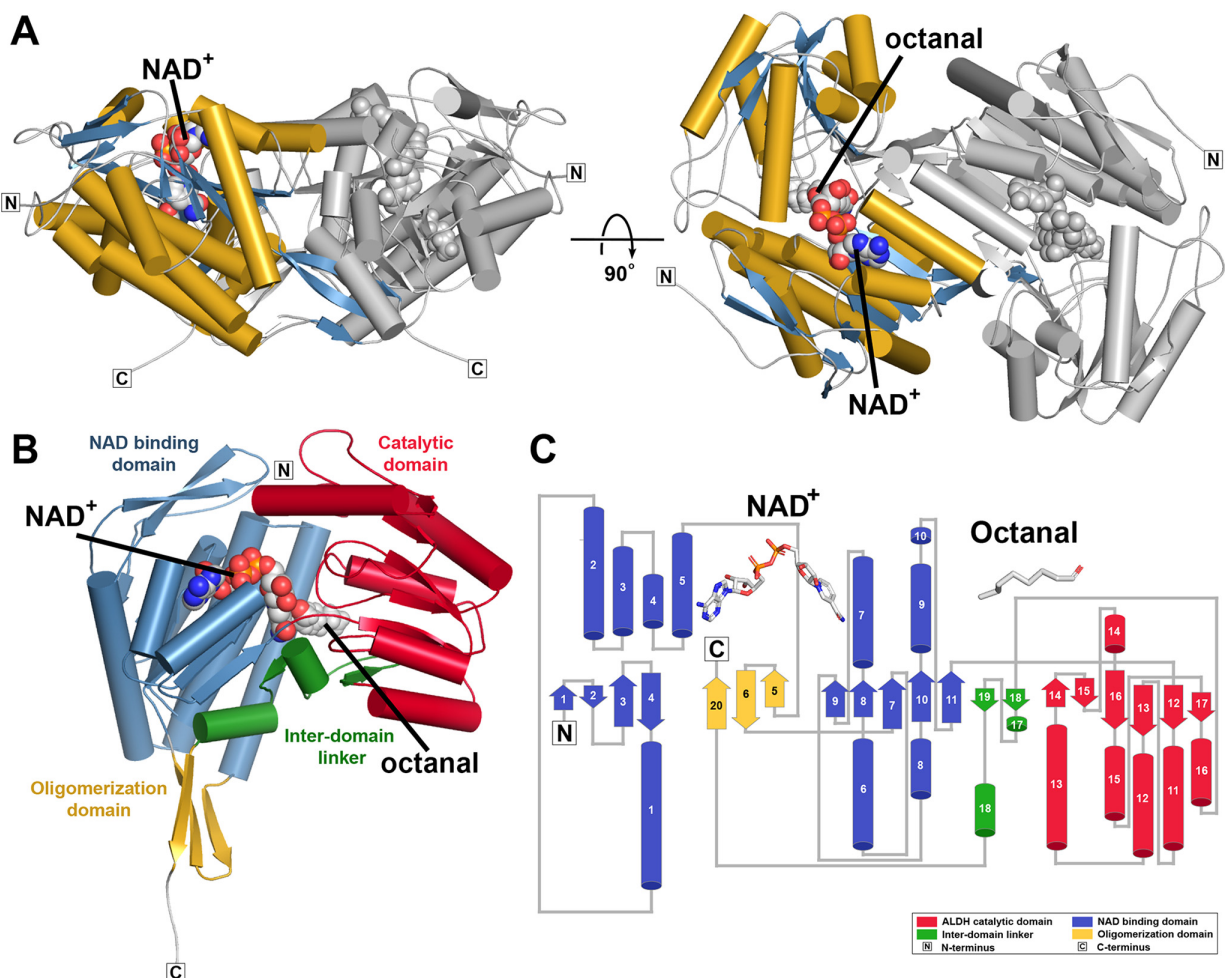


Figure 4. Overall structure of the AldC C291A mutant structure. A, the dimeric structure of AldC dimer is shown as a ribbon diagram. Secondary structure features of monomer A are colored, as follows: α -helices and β -strands are shown as gold cylinders and blue arrows, respectively. The ribbon diagram of monomer B is colored gray with the N and C termini labeled. Ligands are shown as space-filling models. B, domain organization of the AldC monomer. The color scheme for the AldC monomer shows the NAD(H)-binding domains in blue, the catalytic domain in red, the interdomain linker in green, and the oligomerization domain in gold. The positions of octanal and NAD⁺ (space-filling models) are indicated. C, topology of AldC showing the four domains described in B with the same coloring and secondary structure features numbered.

aliphatic substrates (*i.e.* heptanal, octanal, and nonenal), which is consistent with the biochemical assays of AldC. This suggests that substrate preference is governed by van der Waals contacts that maximize nonpolar surface–ligand interactions.

Site-directed mutagenesis of AldC active site residues

To examine the contribution of active site residues, a series of site-directed mutants targeting residues in the NAD(H)-binding site (*i.e.* Trp¹⁵⁸, Lys¹⁸², Glu¹⁸⁵, Thr²³⁴, Ser²³⁶, Glu²⁵⁷, and Glu³⁹¹) and the octanal-binding site (*i.e.* Met¹¹⁴, Leu¹¹⁸, Trp¹⁶⁰, Tyr¹⁶³, Gln¹⁶⁴, Arg²⁸⁵, Ser²⁹², Leu⁴¹⁹, Trp⁴⁵⁰, Phe⁴⁵⁶, and Tyr⁴⁶⁸) were generated. All of the 31 AldC point mutants (M114A, L118A, W158A, W158H, N159A, W160A, W160Y, Y163A, Q164A, Q164N, K182A, K182Q, E185A, E185Q, T234A, T234V, S236A, S236T, E257A, E257Q, R285A, S290A, S292A, E391A, E391D, L419A, L419V, W450A, F456A, F456W, and Y468A) were expressed in *Escherichia*

coli and purified using nickel-affinity and size-exclusion chromatographies. The solubility and purity of each mutant was comparable with WT protein. Under standard assay conditions, enzymatic activities of WT AldC, the AldC C291A mutant (as described above), and the additional 31 point mutants were tested with fixed concentrations of octanal (5 mM) and NAD⁺ (2 mM) using up to 100 μ g of each protein (Fig. 7). The initial enzyme activity screening assays showed that mutation of the catalytic cysteine (C291A) and certain residues in the NAD(H)-binding site (K182A, K182Q, T234V, E257A, and E391A), and octanal-binding site (N159A and L419A) resulted in enzymes with less than 1% of WT specific activity. For example, no activity was detected with the N159A, K182A, and C291A mutants.

Kinetic analysis of mutants with changes in the NAD(H)-binding site of AldC reveals critical residues for biochemical function. As noted above, low activity of the K182A, K182Q, T234V, E257A, and E391A mutants did not allow for accurate

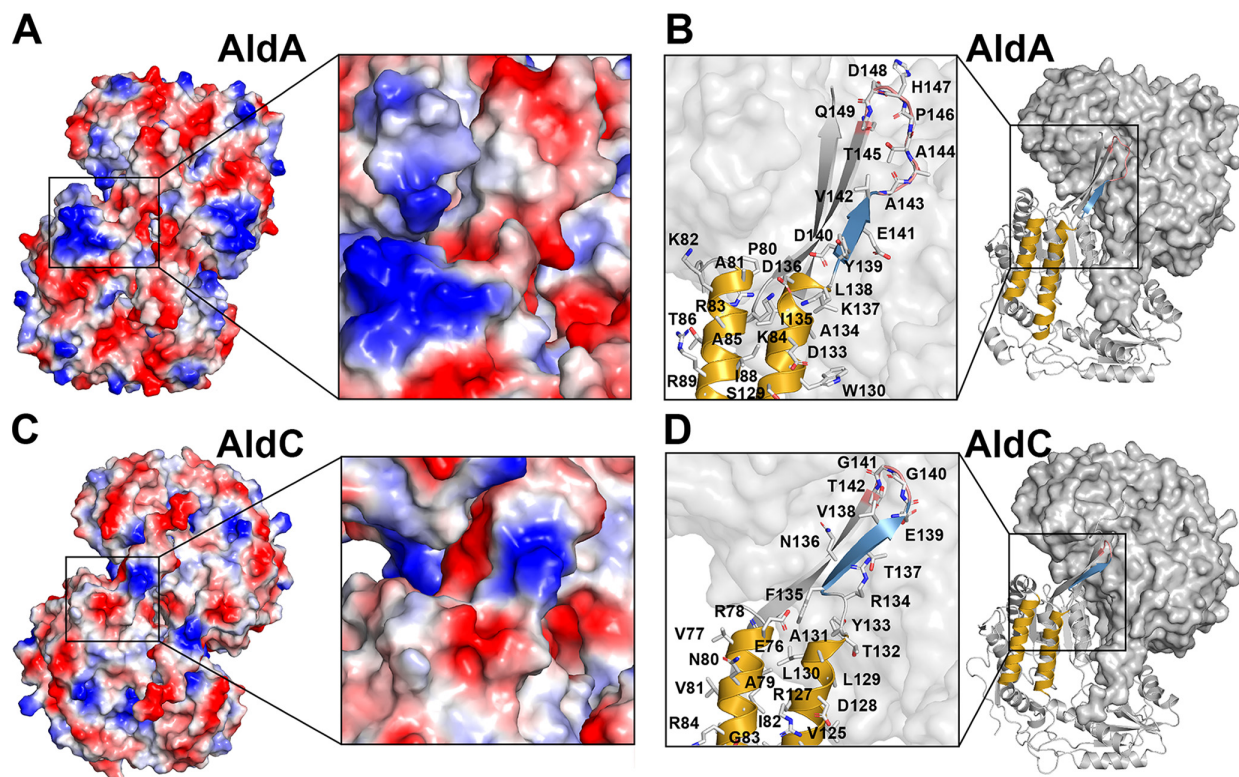


Figure 5. Oligomerization interface differences in AldA (tetramer) and AldC (dimer) from *P. syringae* DC3000. A, electrostatic surface charge of AldA dimer. Electrostatic surface charge was generated using the APBS plugin in PyMOL (Schrödinger) with red (acidic residues) and blue (basic residues). The box corresponds to the oligomerization interface for forming a homotetramer. B, ribbon diagram of monomer A and surface view of monomer B of AldA. The box corresponds to the oligomerization interface for forming a homotetramer. The α -helix and β -sheets for oligomerization are colored blue and gold, respectively. C, electrostatic surface charge of AldC dimer. Coloring is as in A. D, ribbon diagram of monomer A and surface view of monomer B of AldC. Coloring is as in B.

determination of their kinetic parameters (Fig. 7); however, the other mutant proteins with changes in the NAD(H) site were analyzed (Table 3). Mutation of Lys¹⁸² (K182A and K182Q) removes interaction of the lysine side-chain with a hydroxyl group of the adenine ribose of NAD⁺. Similarly, alanine substitutions of Glu²⁵⁷ and Glu³⁹¹ severely disrupted AldC activity (Fig. 7), whereas the E257Q and E391D mutations, which introduced subtler alterations, displayed comparable or less than a 2-fold change in catalytic efficiency (*i.e.* k_{cat}/K_m) with NAD⁺ (Table 3). Loss of the Glu²⁵⁷ side-chain would remove an interaction that helps position the amide group of the nicotinamide and changes at this residue could alter the orientation of the nicotinamide group for hydride transfer (Fig. 6). Similar effects have been observed in other NAD(P)(H)-dependent enzymes (43). Mutation of Glu³⁹¹ to alanine eliminates interaction of the carboxylate group to the nicotinamide ribose group, which can be partially complemented by an aspartate, as seen in the E391D mutant (Table 3). The effects of mutating the third glutamate in the NAD(H)-binding site (*i.e.* Glu¹⁸⁵), which interacts with the adenine ribose group of NAD⁺, are comparable with changes of the other two acidic residues, but not as severe. The E185A mutant displayed a 21-fold decrease in k_{cat}/K_m with NAD⁺, and the E185Q mutant is similar to that of WT AldC (Table 3).

Substitutions of Trp¹⁵⁸, Thr²³⁴, and Ser²³⁶ in the NAD(H)-binding site also reduced activity. The loss of steric bulk with the W158A mutant, which may change how the pyrophosphate backbone of NAD⁺ binds in the site, resulted in a 10-fold decrease in k_{cat}/K_m with NAD⁺, but the W158H mutant modestly reduced efficiency 3-fold (Table 3). Mutation of Thr²³⁴, which is positioned with its methyl group to provide surface contacts to the nicotinamide ring and its hydroxyl group oriented toward the amine of Lys¹⁶⁸ (Fig. 6), led to a slight 5-fold reduction of k_{cat}/K_m with NAD⁺, but nearly inactive protein with a valine mutation (T234V). It is possible that local conformational changes impact the positioning of the nicotinamide ring to impact activity of this mutant protein. The kinetic parameters of S236A and S236T mutations with NAD⁺ were consistent with the loss of a hydrogen bond or changes in the side-chain position for interaction with the pyrophosphate bridge of NAD⁺ (Table 3).

Within the octanal-binding site (Fig. 6), mutations of Asn¹⁵⁹, Trp¹⁶⁰, Ser²⁹², Leu⁴¹⁹, and Phe⁴⁵⁶ were the most disruptive with other changes having more modest effects (Table 3). These residues are located closer to the catalytic site of AldC than those residues with lesser impact on AldC activity. The N159A and L419A mutants lacked significant activity, as noted above. Although the W160A, S292A, L419V, and F456A mutants

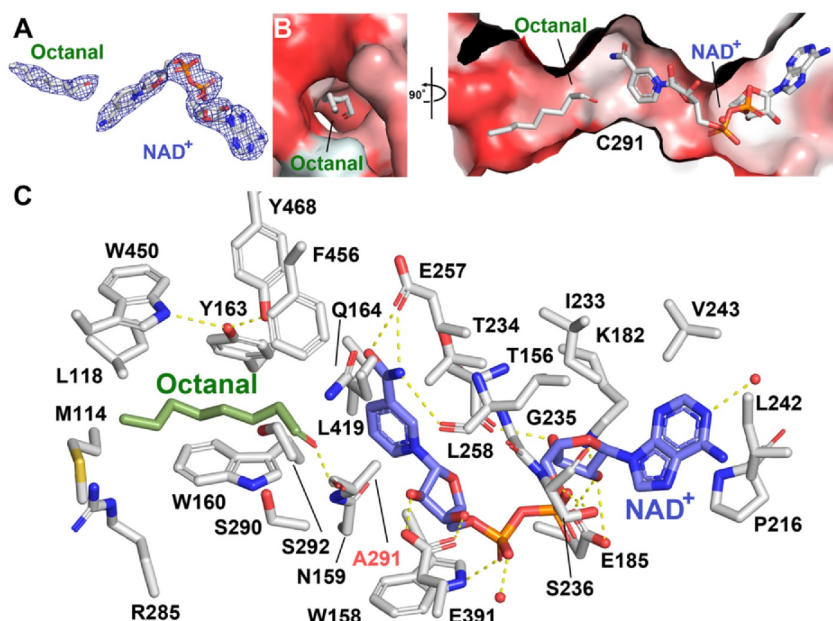


Figure 6. Active site of AldC. A, electron density for octanal (left panels) and NAD⁺ (right panels) in the AldC(C291A)·octanal·NAD⁺ structure is shown as a $2F_o - F_c$ omit map (1.6 σ). B, surface view of the AldC active site. Octanal (substrate) and NAD⁺ (cofactor) are shown as stick models. The active site tunnel is shown as a surface view from the exterior into the catalytic site and the same view rotated 90°. Hydrophobicity was calculated based on the Eisenberg hydrophobicity scale in PyMOL. The darkest red indicates the strongest hydrophobicity, with white indicating the most polar. C, amino acid residues forming interactions with octanal and NAD⁺ in the AldC active site. Side chains of residues interacting with octanal (green) and NAD⁺ (blue) are shown as stick renderings. Two water molecules interacting with the cofactor are shown as red spheres. Hydrogen bond interactions between the amino acid residues and ligands are shown as yellow dotted lines. The C291A mutation is labeled in red.

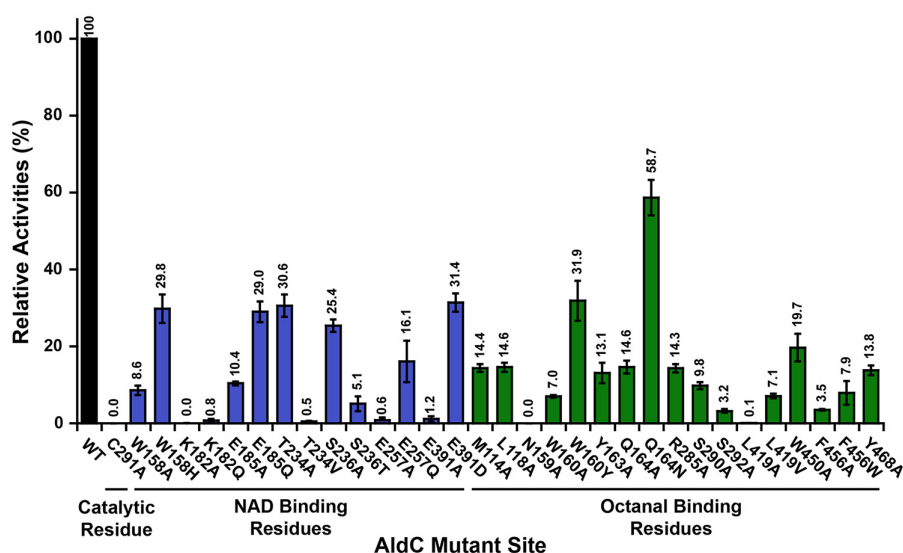


Figure 7. Comparison of WT and mutant AldC specific activities. Mutants proteins are grouped and colored by active site location, as follows: blue, NAD (H)-binding site; and green, octanal-binding site. Standard assay conditions were used as described under "Experimental procedures." Average values $\pm$ S.D. ($n = 3$) are plotted relative to WT specific activity ($1.02 \pm 0.08 \mu\text{mol min}^{-1} \text{mg protein}^{-1}$).

displayed activity in the initial screen experiment (Fig. 7), the estimated K_m values for these proteins with octanal exceeded the solubility limit of this substrate ($\sim 10 \text{ mM}$) in assays, which prevented accurate assessment of kinetic parameters. The N159A mutant abrogates the hydrogen bond between Asn¹⁵⁹ and the carbonyl of octanal; the loss of this interaction likely alters the position of the reactive group for catalysis and results

in the undetected activity of this mutant. Ser²⁹² is also positioned in proximity to the reactive aldehyde group of octanal with mutation to an alanine, resulting in a 30-fold decrease in specific activity (Fig. 7).

Trp¹⁶⁰ provides surface contacts along the central portion of octanal carbon chain with the W160A mutation significantly altering the topology of the binding tunnel. Replacement of

Analysis of AldC from *P. syringae* DC3000

Table 3

Steady-state kinetic analysis of mutant AldC. The protein and location of the mutation are indicated. Enzyme assays were performed as described under "Experimental procedures." Average values $\pm$ S.D. ($n = 3$) are shown.

Protein	Ligand site	Octanal			NAD ⁺		
		K_{cat}	K_m	K_{cat}/K_m	K_{cat}	K_m	K_{cat}/K_m
		min^{-1}	μM	$M^{-1} s^{-1}$	min^{-1}	μM	$M^{-1} s^{-1}$
WT		67.3 ± 1.4	$1,210 \pm 93$	927	30.9 ± 0.4	490 ± 25	1,050
W158A	NAD(H)	2.6 ± 0.1	340 ± 44	127	5.4 ± 0.1	643 ± 57	140
W158H	NAD(H)	13.7 ± 0.5	804 ± 117	284	6.9 ± 0.5	273 ± 38	406
E185A	NAD(H)	4.0 ± 0.2	$4,380 \pm 419$	15	5.2 ± 0.1	$1,900 \pm 120$	46
E185Q	NAD(H)	79.5 ± 2.5	$1,620 \pm 150$	818	15.0 ± 0.2	236 ± 14	1,060
T234A	NAD(H)	21.6 ± 1.2	$1,520 \pm 300$	237	6.8 ± 0.7	594 ± 217	191
S236A	NAD(H)	7.0 ± 0.4	$5,190 \pm 689$	22	6.4 ± 0.2	$1,420 \pm 143$	75
S236T	NAD(H)	3.0 ± 0.2	527 ± 187	95	7.3 ± 0.4	$2,200 \pm 294$	55
E257Q	NAD(H)	9.8 ± 1.0	530 ± 258	308	15.7 ± 4.5	287 ± 170	912
E391D	NAD(H)	25.3 ± 2.4	$2,730 \pm 743$	154	10.3 ± 3.0	248 ± 183	692
M114A	Octanal	22.0 ± 1.0	$1,220 \pm 218$	301	11.7 ± 0.5	543 ± 80	360
L118A	Octanal	17.0 ± 1.7	$1,080 \pm 432$	262	11.0 ± 0.4	298 ± 34	615
W160Y	Octanal	63.7 ± 4.5	$2,160 \pm 472$	491	6.1 ± 0.6	516 ± 26	197
Y163A	Octanal	70.2 ± 8.7	$6,540 \pm 1,580$	179	18.8 ± 5.3	244 ± 162	1,280
Q164A	Octanal	76.3 ± 17.6	$4,110 \pm 2,250$	309	29.2 ± 5.3	416 ± 213	1,170
Q164N	Octanal	66.0 ± 7.9	$2,270 \pm 819$	485	24.8 ± 1.5	258 ± 38	1,600
R285A	Octanal	16.5 ± 0.5	$1,100 \pm 128$	250	11.8 ± 0.5	434 ± 66	453
S290A	Octanal	2.8 ± 0.1	830 ± 180	56	7.5 ± 0.2	$1,220 \pm 81$	102
W450A	Octanal	9.8 ± 1.0	220 ± 90	742	5.1 ± 0.3	805 ± 167	106
F456W	Octanal	140 ± 33	$5,420 \pm 2,680$	431	2.9 ± 0.1	622 ± 5	78
Y468A	Octanal	8.4 ± 0.6	$2,520 \pm 548$	56	8.6 ± 0.3	743 ± 81	193

Trp¹⁶⁰ with a tyrosine resulted a less than 2-fold change in catalytic efficiency (Table 3). Leu⁴¹⁹ is also positioned at the intersection of the NAD(H) and octanal-binding sites and forms van der Waals contacts with Ala²⁹¹ in the AldC structure. Changes of this residue to either alanine (L419A) or valine (L419V) may alter the orientation of the catalytic sulfhydryl required for the conversion of octanal to its carboxylic acid. In the AldC crystal structure, Phe⁴⁵⁶ π -stacks with Tyr⁴⁶⁸, which forms an interaction network with Tyr¹⁶³ and Trp⁴⁵⁰ (Fig. 6). Mutation of Phe⁴⁵⁶ to an alanine (F456A) also increases the apparent K_m of octanal, as described above, which likely results from altered packing along the aldehyde-binding site of the F456A mutant. In contrast, the F456W mutant results in a modest 2-fold decrease in k_{cat}/K_m with octanal (Table 3).

The opposites of Trp¹⁶⁰, Tyr⁴⁶⁸, Tyr¹⁶³, and Trp⁴⁵⁰ are interconnected through a series of hydrogen bonds to form one of the walls of the octanal-binding site (Fig. 6). Disruptions in this network using the Y468A, Y163A, and W450A mutants resulted in 20-, 5-, and 9-fold decreases in catalytic efficiency with octanal (Table 3). It appears that removal on one of these aromatic side chains alters AldC activity but is not sufficient to significantly disrupt substrate binding and catalysis. Alanine substitutions of the three residues situated toward the solvent entrance of the aldehyde-binding site, Met¹¹⁴, Leu¹¹⁸, and Arg²⁸⁵, have only slight 3–4-fold reductions in catalytic efficiency compared with WT AldC (Table 3), which is consistent with the interior hydrophobic surface as the major determinant of substrate interaction. Two other residues in the octanal-binding site, Gln¹⁶⁴ and Ser²⁹⁰, also play lesser roles in substrate binding, as suggested by the modest effects of the Q164A and Q164N mutations and the S290A mutant (Table 3).

Conclusion

The genus *Pseudomonas* is one of the most ubiquitous and complex among the Gram-negative bacteria (22). Because many *Pseudomonas* species evolved to grow under unfavorable

environmental conditions (*i.e.* severe nutrient limitation, extreme temperatures, high salinity, and low oxygen or water availability), they also evolved metabolic diversity and plasticity to use a variety of nutrient sources (*i.e.* carbon, nitrogen, and sulfur), to detoxify toxic organic chemicals, and to produce multiple specialized metabolites, including polymers and small molecule compounds (31, 44). In particular, *P. syringae*, a species that includes many plant pathogenic strains, developed diverse bacterial virulence mechanisms to survive in the adverse environmental conditions of the phyllosphere (45). Further, recent studies suggest that *Pto*DC3000 uses several strategies to manipulate the auxin biology in its host plants to promote pathogenicity, including using an indole-3-acetaldehyde dehydrogenase for IAA synthesis (21). Recently, a genus-wide study identified a total of 6,510 aldehyde dehydrogenase sequences in 258 *Pseudomonas* strains belonging to 46 different species, but only a limited number of newly identified *Pseudomonas* aldehyde dehydrogenases have been biochemically characterized (32).

In contrast to the better-studied aromatic aldehyde dehydrogenases identified from various *Pseudomonas* species (46–48), which have been studied because of their ability to catabolize and biodegrade a range of compounds, including naphthalene-derived molecules, and metabolic aldehyde dehydrogenases linked to the osmoprotectant betaine and growth in biofilms (49–51), less is understood about the biochemistry of aliphatic aldehyde dehydrogenases in these microbes. Here we comprehensively examined and identified AldC from *Pto*DC3000 as a long-chain aliphatic aldehyde dehydrogenase using an integrated approach combining phylogenetic, crystallographic, and biochemical analyses.

The biochemical activity of AldC from *Pto*DC3000 is consistent with the traditional role of aldehyde dehydrogenases as metabolic clean-up enzymes that convert reactive aldehydes into less active carboxylates (11, 12); however, growing evidence suggests that these enzymes may have broader

contributions to the specialized environments that different *Pseudomonas* species occupy. For example, recent work identified AldA as an indole-3-acetaldehyde dehydrogenase involved in the synthesis of the phytohormone IAA in the plant pathogen *P. syringae* DC3000 and demonstrated a role for this protein in virulence of the pathogen in plant leaves (21). Given the growth environment of this pathogenic *Pseudomonas* species in the interior of plant leaves, it is possible that aliphatic molecules in the apoplast could contribute to energy sources for the microbe or to the synthesis of surfactants that help establish microbial colonization in the host plant (52–54). The physiological contribution of AldC to the life cycle of *Pto*DC3000 and related plant pathogenic or plant-associated species and strains remains to be evaluated but raises the question of how these microbes develop host relationships in the natural environment. Overall, this study provides molecular insight for understanding the evolution of the prokaryotic aldehyde dehydrogenase superfamily and for the potential development of inhibitors for the pathogenic *Pseudomonas* species, which could be useful for pathogen control in agriculture.

Experimental procedures

Cloning and site-directed mutagenesis

The pET28a-AldC construct was previously described (21). Briefly, the full-length coding sequence of AldC (UniProt: PSPTO_3644; GenBankTM: WP_011104646.1) was amplified from *P. syringae* strain *Pto*DC3000 genomic DNA and cloned into the pBlunt II-TOPO vector (Invitrogen). The AldC fragment from the pBlunt II-TOPO-AldC was excised using *Nde*I and *Xho*I and subcloned into pET-28a (Novagen). Site-directed mutants of AldC (M114A, L118A, W158A, W158H, N159A, W160A, W160Y, Y163A, Q164A, Q164N, K182A, K182Q, E185A, E185Q, T234A, T234V, S236A, S236T, E257A, E257Q, R285A, S290A, C291A, S292A, E391A, E391D, L419A, L419V, W450A, F456A, F456W, and Y468A) were generated using the QuikChange PCR method (Agilent Technologies) with the pET28a-AldC construct as template. For protein expression, each pET28a-AldC construct was transformed into *E. coli* BL21 (DE3) (EMD Millipore).

Protein expression and purification

Transformed *E. coli* BL21(DE3) cells containing either WT or mutant AldC construct were grown at 37 °C in Terrific broth with 50 $\mu\text{g ml}^{-1}$ kanamycin until $A_{600\text{ nm}} = \sim 0.6\text{--}0.9$. After induction with 1 mM isopropyl-1-thio- β -D-galactopyranoside, the cells were grown at 18 °C overnight. Following centrifugation (5,000 $\times g$ for 30 min), the cell pellets were resuspended in 50 mM Tris, pH 8.0, 500 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, and 1% (v/v) Tween 20. Following lysis by sonication, the cell debris was removed by centrifugation (12,000 $\times g$ for 45 min), and the supernatant was loaded onto a Ni^{2+} -nitriloacetic acid column (Qiagen). The column was washed with 50 mM Tris, pH 8.0, 500 mM NaCl, 25 mM imidazole, and 10% (v/v) glycerol to remove unbound proteins. Column-bound AldC protein was eluted using 50 mM Tris, pH 8.0, 500 mM NaCl, 25 mM imidazole, 10% glycerol, and 250 mM

imidazole. AldC was further purified by size-exclusion chromatography with a Superdex 200 16/60 size-exclusion column (GE Healthcare) equilibrated in 25 mM Hepes (pH 7.5) and 100 mM NaCl. Fractions corresponding to the purified AldC protein were pooled and concentrated to $\sim 9\text{ mg ml}^{-1}$. Protein concentrations were determined using the Bradford method with BSA as a standard.

Enzyme assays and steady-state kinetic analysis

Enzymatic activity of WT and mutant AldC was measured by monitoring NADH formation ($\epsilon_{340\text{ nm}} = 6220\text{ M}^{-1}\text{ cm}^{-1}$; 100- μl volume) at $A_{340\text{ nm}}$ using an EPOCH2 microplate spectrophotometer (BioTek). The substrate and cofactor screening experiments were performed at 25 °C in a standard assay mix of 100 mM Tris-HCl (pH 8.0) and 100 mM KCl with a fixed concentration of NAD^+ (2 mM) and aldehyde (5 mM). The panel of aldehydes and their chemical structures used for substrate screening is shown in Table S2. Steady-state kinetic parameters of WT and mutant AldC were determined at 25 °C in a standard assay mix with either fixed NAD^+ (2.0 mM) and varied aldehyde (*i.e.* valeraldehyde, hexanal, heptanal, octanal, nonanal, hydrocinnamaldehyde, or 4-pyridinecarboxaldehyde; 0.01–10 mM) or with fixed octanal (5 mM) and varied cofactor (NAD^+ or NADP^+ ; 0.05–5 mM). Protein amounts ranged from 0.1 μg for WT AldC to 100 μg for less active AldC mutants. The resulting initial velocity data were fit to the Michaelis–Menten equation, $v = (k_{\text{cat}}[\text{S}])/(K_m + [\text{S}])$, using SigmaPlot.

Protein crystallography

Protein crystals of the AldC C291A mutant were grown by the hanging-drop vapor-diffusion method at 4 °C. Crystals of the AldC C291A mutant (9 mg ml^{-1}) in complex with octanal and NAD^+ formed grew in drops of a 1:1 mixture of proteins and crystallization buffer: 20% (w/v) PEG-1000, 100 mM Tris-HCl (pH 7.0), 2 mM octanal, and 5 mM NAD^+ . Crystals were stabilized in cryoprotectant (mother liquor with either 30% (v/v) glycerol or 30% (v/v) ethylene glycol) before flash-freezing in liquid nitrogen for data collection at 100 K. Diffraction data were collected at Beamline 19-ID of the Advanced Photon Source at the Argonne National Laboratory HKL3000 was used to index, integrate, and scale the collected X-ray data (55). Molecular replacement was used to solve the X-ray crystal structure of AldC using PHASER (56) with the three-dimensional structure of the AldA indole-3-acetaldehyde dehydrogenase from *P. syringae*, which shares 39% amino acid identity with AldC, as a search model (Protein Data Bank code 5IUW) (21). COOT (57) and PHENIX (58) were used for iterative rounds of model building and refinement, respectively. Data collection and refinement statistics are summarized in Table 2. The final model included $\text{Leu}^6\text{--Leu}^{490}$ of chain A, $\text{Leu}^6\text{--Asp}^{489}$ of chain B, 1 octanal, and 1 NAD^+ in each chain, and 197 waters. Atomic coordinates and structure factors for the AldC(C291A)·octanal· NAD^+ complex were deposited in the RCSB Protein Data Bank (code 6X9L).

Analysis of AldC from *P. syringae* DC3000

Data availability

All data are contained in the article with structural information deposited in the Protein Data Bank (code [6X9L](#)), as noted above.

Author contributions—S. G. L., K. H., S. A. M., B. N. K., and J. M. J. conceptualization; S. G. L. data curation; S. G. L., K. H., and W. C. formal analysis; S. G. L., C. K. H., B. C., W. C., B. N. K., and J. M. J. supervision; S. G. L., K. H., O. A., S. O. A., A. G. B., H.-Y. S. C., M. D., C. M. F., A. R. G., J. K., J. S. K., M. C. L., J. J. L., R. L., J. E. L., C. M. M., A. F. M., S. H. P., V. P., C. D. R., A. S., S. S., J. D. T., B. R. C., C. K. H., B. C., and W. C. investigation; S. G. L., K. H., O. A., S. O. A., A. G. B., H.-Y. S. C., M. D., C. M. F., A. R. G., J. K., J. S. K., M. C. L., J. J. L., R. L., J. E. L., C. M. M., A. F. M., S. H. P., V. P., C. D. R., A. S., S. S., J. D. T., B. R. C., C. K. H., B. C., and W. C. methodology; S. G. L., K. H., S. A. M., B. N. K., and J. M. J. writing-original draft; S. G. L., K. H., S. A. M., B. N. K., and J. M. J. writing-review and editing; S. A. M., B. N. K., and J. M. J. resources; S. A. M. and J. M. J. validation; B. N. K. and J. M. J. funding acquisition; B. N. K. and J. M. J. project administration.

Funding and additional information—This work was supported by National Science Foundation Grants IOS-1645908 (to B. N. K.) and MCB-1614539 (to J. M. J.) and a grant from the Howard Hughes Medical Institute Science Education Program (to J. M. J.). C.K.H. was supported by National Science Foundation Graduate Research Fellowship DGE-1143954. This project was part of Bio4522, an upper-level laboratory course at Washington University in St. Louis. Portions of this research were carried out at the Argonne National Laboratory Structural Biology Center of the Advanced Photon Source, a national user facility operated by the University of Chicago for the Department of Energy Office of Biological and Environmental Research under Grant DE-AC02-06CH11357.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: IAA, indole-3-acetic acid; RMSD, root-mean-square deviation.

References

1. Michael, A. J. (2017) Evolution of biosynthetic diversity. *Biochem. J.* **474**, 2277–2299 [CrossRef Medline](#)
2. Newton, M. S., Arcus, V. L., Gerth, M. L., and Patrick, W. M. (2018) Enzyme evolution: innovation is easy, optimization is complicated. *Curr. Opin. Struct. Biol.* **48**, 110–116 [CrossRef Medline](#)
3. Guengerich, F. P. (2008) Cytochrome P450 and chemical toxicology. *Chem. Res. Toxicol.* **21**, 70–83 [CrossRef Medline](#)
4. Nelson, D. R. (1999) Cytochrome P450 and the individuality of species. *Arch. Biochem. Biophys.* **369**, 1–10 [CrossRef Medline](#)
5. Persson, B., Hedlund, J., and Jörnvall, H. (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families: the MDR superfamily. *Cell. Mol. Life Sci.* **65**, 3879–3894 [CrossRef Medline](#)
6. Jörnvall, H., Persson, B., Krook, M., Atrian, S., González-Duarte, R., Jeffery, J., and Ghosh, D. (1995) Short-chain dehydrogenases/reductases (SDR). *Biochemistry* **34**, 6003–6013 [CrossRef Medline](#)
7. Penning, T. M. (2015) The aldo-keto reductases (AKRs): overview. *Chem.-Biol. Interact.* **234**, 236–246 [CrossRef Medline](#)
8. Jez, J. M., Bennett, M. J., Schlegel, B. P., Lewis, M., and Penning, T. M. (1997) Comparative anatomy of the aldo-keto reductase superfamily. *Biochem. J.* **326**, 625–636 [CrossRef Medline](#)
9. Yoshida, A., Rzhetsky, A., Hsu, L. C., and Chang, C. (1998) Human aldehyde dehydrogenase gene family. *Eur. J. Biochem.* **251**, 549–557 [CrossRef Medline](#)
10. Brocker, C., Vasiliou, M., Carpenter, S., Carpenter, C., Zhang, Y., Wang, X., Kotchoni, S. O., Wood, A. J., Kirch, H. H., Kopečný, D., Nebert, D. W., and Vasiliou, V. (2013) Aldehyde dehydrogenase (ALDH) superfamily in plants: gene nomenclature and comparative genomics. *Planta* **237**, 189–210 [CrossRef Medline](#)
11. O'Brien, P. J., Siraki, A. G., and Shangari, N. (2005) Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health. *Crit. Rev. Toxicol.* **35**, 609–662 [CrossRef Medline](#)
12. Vasiliou, V., Pappa, A., and Estey, T. (2004) Role of human aldehyde dehydrogenases in endobiotic and xenobiotic metabolism. *Drug Metab. Rev.* **36**, 279–299 [CrossRef Medline](#)
13. Klyosov, A. A. (1996) Kinetics and specificity of human liver aldehyde dehydrogenases toward aliphatic, aromatic, and fused polycyclic aldehydes. *Biochemistry* **35**, 4457–4467 [CrossRef Medline](#)
14. Steinmetz, C. G., Xie, P., Weiner, H., and Hurley, T. D. (1997) Structure of mitochondrial aldehyde dehydrogenase: the genetic component of ethanol aversion. *Structure* **5**, 701–711 [CrossRef Medline](#)
15. Ni, L., Zhou, J., Hurley, T. D., and Weiner, H. (1999) Human liver mitochondrial aldehyde dehydrogenase: three-dimensional structure and the restoration of solubility and activity of chimeric forms. *Protein Sci.* **8**, 2784–2790 [CrossRef Medline](#)
16. Brocker, C., Lassen, N., Estey, T., Pappa, A., Cantore, M., Orlova, V., Chavakis, T., Kavanagh, K. L., Oppermann, U., and Vasiliou, V. (2010) Aldehyde dehydrogenase 7A1 (ALDH7A1) is a novel enzyme involved in cellular defense against hyperosmotic stress. *J. Biol. Chem.* **285**, 18452–18463 [CrossRef Medline](#)
17. Nair, R. B., Bastress, K. L., Ruegger, M. O., Denault, J. W., and Chapple, C. (2004) The *Arabidopsis thaliana* REDUCED EPIDERMAL FLUORESCENCE1 gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis. *Plant Cell* **16**, 544–554 [CrossRef Medline](#)
18. Bosch, M., Mayer, C. D., Cookson, A., and Donnison, I. S. (2011) Identification of genes involved in cell wall biogenesis in grasses by differential gene expression profiling of elongating and non-elongating maize internodes. *J. Exp. Bot.* **62**, 3545–3561 [CrossRef Medline](#)
19. Kotchoni, S. O., Kuhns, C., Ditzler, A., Kirch, H. H., and Bartels, D. (2006) Over-expression of different aldehyde dehydrogenase genes in *Arabidopsis thaliana* confers tolerance to abiotic stress and protects plants against lipid peroxidation and oxidative stress. *Plant Cell Environ.* **29**, 1033–1048 [CrossRef Medline](#)
20. Rodrigues, S. M., Andrade, M. O., Gomes, A. P., Damatta, F. M., Barakat-Pereira, M. C., and Fontes, E. P. (2006) *Arabidopsis* and tobacco plants ectopically expressing the soybean antiquitin-like ALDH7 gene display enhanced tolerance to drought, salinity, and oxidative stress. *J. Exp. Bot.* **57**, 1909–1918 [CrossRef Medline](#)
21. McClerklin, S. A., Lee, S. G., Harper, C. P., Nwumeh, R., Jez, J. M., and Kunkel, B. N. (2018) Indole-3-acetaldehyde dehydrogenase-dependent auxin synthesis contributes to virulence of *Pseudomonas syringae* strain DC 3000. *PLoS Pathog.* **14**, e1006811 [CrossRef Medline](#)
22. Xin, X. F., and He, S. Y. (2013) *Pseudomonas syringae* pv. *tomato* DC3000: a model pathogen for probing disease susceptibility and hormone signaling in plants. *Annu. Rev. Phytopathol.* **51**, 473–498 [CrossRef Medline](#)
23. Jones, J. D. G., and Dangl, J. L. (2006) The plant immune system. *Nature* **444**, 323–329 [CrossRef Medline](#)
24. Chisholm, S. T., Coaker, G., Day, B., and Staskawicz, B. J. (2006) Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* **124**, 803–814 [CrossRef Medline](#)

25. Djami-Tchatchou, A. T., Harrison, G. A., Harper, C. P., Wang, R., Prigge, M., Estelle, M., and Kunkel, B. N. (2020) Dual role of auxin in regulating plant defense and bacterial virulence gene expression during *Pseudomonas syringae* Pto DC3000 pathogenesis. *Mol. Plant Microbe Interact.* **33**, 1059–1071 [CrossRef Medline](#)
26. Kunkel, B. N., and Harper, C. P. (2018) The roles of auxin during interactions between bacterial plant pathogens and their hosts. *J. Exp. Bot.* **69**, 245–254 [CrossRef Medline](#)
27. Aragón, I. M., Pérez-Martínez, I., Moreno-Pérez, A., Cerezo, M., and Ramos, C. (2014) New insights into the role of indole-3-acetic acid in the virulence of *Pseudomonas savastanoi* pv. *savastanoi*. *FEMS Microbiol. Lett.* **356**, 184–192 [CrossRef Medline](#)
28. Mutka, A. M., Fawley, S., Tsao, T., and Kunkel, B. N. (2013) Auxin promotes susceptibility to *Pseudomonas syringae* via a mechanism independent of suppression of salicylic acid-mediated defenses. *Plant J.* **74**, 746–754 [CrossRef Medline](#)
29. Manulis, S., Haviv-Chesner, A., Brandl, M. T., Lindow, S. E., and Barash, I. (1998) Differential involvement of indole-3-acetic acid biosynthetic pathways in pathogenicity and epiphytic fitness of *Erwinia herbicola* pv. *gypsophylae*. *Mol. Plant Microbe Interact.* **11**, 634–642 [CrossRef Medline](#)
30. Xie, B., Xu, K., Zhao, H. X., and Chen, S. F. (2005) Isolation of transposon mutants from *Azospirillum brasilense* Yu62 and characterization of genes involved in indole-3-acetic acid biosynthesis. *FEMS Microbiol. Lett.* **248**, 57–63 [CrossRef Medline](#)
31. Sophos, N. A., and Vasiliou, V. (2003) Aldehyde dehydrogenase gene superfamily: the 2002 update. *Chem. Biol. Interact.* **143–144**, 5–22 [CrossRef Medline](#)
32. Riveros-Rosas, H., Julián-Sánchez, A., Moreno-Hagelsieb, G., and Muñoz-Clares, R. A. (2019) Aldehyde dehydrogenase diversity in bacteria of the *Pseudomonas* genus. *Chem. Biol. Interact.* **304**, 83–87 [CrossRef Medline](#)
33. Lamb, A. L., and Newcomer, M. E. (1999) The structure of retinal dehydrogenase type II at 2.7 Å resolution: implications for retinal specificity. *Biochemistry* **38**, 6003–6011 [CrossRef Medline](#)
34. Bateman, O. A., Purkiss, A. G., van Montfort, R., Slingsby, C., Graham, C., and Wistow, G. (2003) Crystal structure of eta-crystallin: adaptation of a class 1 aldehyde dehydrogenase for a new role in the eye lens. *Biochemistry* **42**, 4349–4356 [CrossRef Medline](#)
35. Díaz-Sánchez, A. G., González-Segura, L., Mújica-Jiménez, C., Rudiño-Piñera, E., Montiel, C., Martínez-Castilla, L. P., and Muñoz-Clares, R. A. (2012) Amino acid residues critical for the specificity for betaine aldehyde of the plant ALDH10 isoenzyme involved in the synthesis of glycine betaine. *Plant Physiol.* **158**, 1570–1582 [CrossRef Medline](#)
36. Končítiková, R., Vigouroux, A., Kopečná, M., Andree, T., Bartoš, J., Šebela, M., Moréra, S., and Kopečný, D. (2015) Role and structural characterization of plant aldehyde dehydrogenases from family 2 and family 7. *Biochem. J.* **468**, 109–123 [CrossRef Medline](#)
37. Koch, M. F., Harteis, S., Blank, I. D., Pestel, G., Tietze, L. F., Ochsenfeld, C., Schneider, S., and Sieber, S. A. (2015) Structural, biochemical, and computational studies reveal the mechanism of selective aldehyde dehydrogenase 1A1 inhibition by cytotoxic duocarmycin analogues. *Angew. Chem. Int. Ed. Engl.* **54**, 13550–13554 [CrossRef Medline](#)
38. Korasick, D. A., White, T. A., Chakravarthy, S., and Tanner, J. J. (2018) NAD⁺ promotes assembly of the active tetramer of aldehyde dehydrogenase 7A1. *FEBS Lett.* **592**, 3229–3238 [CrossRef Medline](#)
39. González-Segura, L., Riveros-Rosas, H., Julián-Sánchez, A., and Muñoz-Clares, R. A. (2015) Residues that influence coenzyme preference in the aldehyde dehydrogenases. *Chem. Biol. Interact.* **234**, 59–74 [CrossRef Medline](#)
40. Riveros-Rosas, H., González-Segura, L., Julián-Sánchez, A., Díaz-Sánchez, A. G., and Muñoz-Clares, R. A. (2013) Structural determinants of substrate specificity in aldehyde dehydrogenases. *Chem. Biol. Interact.* **202**, 51–61 [CrossRef Medline](#)
41. Luo, M., and Tanner, J. J. (2015) Structural basis of substrate recognition by aldehyde dehydrogenase 7A1. *Biochemistry* **54**, 5513–5522 [CrossRef Medline](#)
42. Pemberton, T. A., and Tanner, J. J. (2013) Structural basis of substrate selectivity of $\Delta(1)$ -pyrroline-5-carboxylate dehydrogenase (ALDH4A1): semialdehyde chain length. *Arch. Biochem. Biophys.* **538**, 34–40 [CrossRef Medline](#)
43. Ma, H., Ratnam, K., and Penning, T. M. (2000) Mutation of nicotinamide pocket residues in rat liver 3 α -hydroxysteroid dehydrogenase reveals different modes of cofactor binding. *Biochemistry* **39**, 102–109 [CrossRef Medline](#)
44. Gomila, M., Peña, A., Mulet, M., Lalucat, J., and García-Valdés, E. (2015) Phylogenomics and systematics in *Pseudomonas*. *Front. Microbiol.* **6**, 214 [CrossRef Medline](#)
45. Lindow, S. E., and Leveau, J. H. (2002) Phyllosphere microbiology. *Curr. Opin. Biotechnol.* **13**, 238–243 [CrossRef Medline](#)
46. Coitinho, J. B., Pereira, M. S., Costa, D. M., Guimarães, S. L., Araújo, S. S., Hengge, A. C., Brandão, T. A., and Nagem, R. A. (2016) Structural and kinetic properties of the aldehyde dehydrogenase NahF, a broad substrate specificity enzyme for aldehyde oxidation. *Biochemistry* **55**, 5453–5463 [CrossRef Medline](#)
47. Crabo, A. G., Singh, B., Nguyen, T., Emami, S., Gassner, G. T., and Sazinsky, M. H. (2017) Structure and biochemistry of phenylacetaldehyde dehydrogenase from the *Pseudomonas putida* S12 styrene catabolic pathway. *Arch. Biochem. Biophys.* **616**, 47–58 [CrossRef Medline](#)
48. Zahniser, M. P. D., Prasad, S., Kneen, M. M., Kreinbring, C. A., Petsko, G. A., Ringe, D., and McLeish, M. J. (2017) Structure and mechanism of benzaldehyde dehydrogenase from *Pseudomonas putida* ATCC 12633, a member of the class 3 aldehyde dehydrogenase superfamily. *Protein Eng. Des. Sel.* **30**, 271–278 [CrossRef Medline](#)
49. Wargo, M. J. (2013) Choline catabolism to glycine betaine contributes to *Pseudomonas aeruginosa* survival during murine lung infection. *PLoS One* **8**, e56850 [CrossRef Medline](#)
50. Muñoz-Clares, R. A., Díaz-Sánchez, A. G., González-Segura, L., and Montiel, C. (2010) Kinetic and structural features of betaine aldehyde dehydrogenases: mechanistic and regulatory implications. *Arch. Biochem. Biophys.* **493**, 71–81 [CrossRef Medline](#)
51. Velasco-García, R., Chacón-Aguilar, V. M., Hervert-Hernández, D., and Muñoz-Clares, R. A. (2003) Inactivation of betaine aldehyde dehydrogenase from *Pseudomonas aeruginosa* and *Amaranthus hypochondriacus* L. leaves by disulfiram. *Chem. Biol. Interact.* **143–144**, 149–158 [CrossRef Medline](#)
52. Carnicero, D., Fernández-Valverde, M., Cañedo, L. M., Schleissner, C., and Luengo, J. M. (1997) Octanoic acid uptake in *Pseudomonas putida* U. *FEMS Microbiol. Lett.* **149**, 51–58 [CrossRef](#)
53. Rosenblatt, J., Reitzel, R. A., Vargus-Cruz, N., Chafari, A. M., Hachem, R., and Raad, I. (2017) Caprylic and polygalacturonic acid combinations for eradication of microbial organism embedded in biofilm. *Front. Micro.* **8**, 1999
54. Oso, S., Walters, M., Schlechter, R. O., and Remus-Emsermann, M. N. P. (2019) Utilisation of hydrocarbons and production of surfactants by bacterial isolated from plant leaf surfaces. *FEMS Microbiol. Lett.* **366**, fnz061 [Medline](#)
55. Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution: from diffraction images to an initial model in minutes. *Acta Crystallogr. D* **62**, 859–866 [CrossRef Medline](#)
56. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 [CrossRef Medline](#)
57. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 [CrossRef Medline](#)
58. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., et al. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 [CrossRef Medline](#)

Analysis of AldC from *P. syringae* DC3000

59. Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S. Y., *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 [CrossRef Medline](#)
60. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 [CrossRef Medline](#)
61. Kumar, S., Stecher, G., and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 [CrossRef Medline](#)
62. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282 [CrossRef Medline](#)