PORTLAND PRESS

Research Article

Investigating the reaction and substrate preference of indole-3-acetaldehyde dehydrogenase from the plant pathogen *Pseudomonas syringae Pto*DC3000

Kaleena Zhang[1,*], Josephine S. Lee[1,*], Regina Liu[1,*], Zita T. Chan[1], Trenton J. Dawson[1], Elisa S. De Togni[1], Chris T. Edwards[1], Isabel K. Eng[1], Ashley R. Gao[1], Luis A. Goicouria[1], Erin M. Hall[1], Kelly A. Hu[1], Katherine Huang[1], Alexander Kizhner[1], Kelsie C. Kodama[1], Andrew Z. Lin[1], Jennifer Y. Liu[1], Alan Y. Lu[1], Owen W. Peng[1], Erica P. Ryu[1], Sophia Shi[1], Maria L. Sorkin[1], Patricia L. Walker[1], Grace J. Wang[1], Mark C. Xu[1], Rebecca S. Yang[1], Barrie Cascella[1], Wilhelm Cruz[1], Cynthia K. Holland[1,2], Sheri A. McClerkin[1,3], Barbara N. Kunkel[1], Soon Goo Lee[1,4] and Joseph M. Jez[1]

[1]Department of Biology, Washington University in St. Louis, St. Louis, MO 63130, U.S.A.; [2]Department of Biology, Williams College, Williamstown, MA 01267, U.S.A.; [3]Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637, U.S.A.; [4]Department of Chemistry and Biochemistry, University of North Carolina Wilmington, Wilmington, NC 28403, U.S.A.

Correspondence: Joseph M. Jez (jjez@wustl.edu)

OPEN ACCESS

Aldehyde dehydrogenases (ALDHs) catalyze the conversion of various aliphatic and aromatic aldehydes into corresponding carboxylic acids. Traditionally considered as housekeeping enzymes, new biochemical roles are being identified for members of ALDH family. Recent work showed that AldA from the plant pathogen *Pseudomonas syringae* strain *Pto*DC3000 (*Pto*DC3000) functions as an indole-3-acetaldehyde dehydrogenase for the synthesis of indole-3-acetic acid (IAA). IAA produced by AldA allows the pathogen to suppress salicylic acid-mediated defenses in the model plant *Arabidopsis thaliana*. Here we present a biochemical and structural analysis of the AldA indole-3-acetaldehyde dehydrogenase from *Pto*DC3000. Site-directed mutants targeting the catalytic residues Cys[302] and Glu[267] resulted in a loss of enzymatic activity. The X-ray crystal structure of the catalytically inactive AldA C302A mutant in complex with IAA and NAD$^+$ showed the cofactor adopting a conformation that differs from the previously reported structure of AldA. These structures suggest that NAD$^+$ undergoes a conformational change during the AldA reaction mechanism similar to that reported for human ALDH. Site-directed mutagenesis of the IAA binding site indicates that changes in the active site surface reduces AldA activity; however, substitution of Phe[169] with a tryptophan altered the substrate selectivity of the mutant to prefer octanal. The present study highlights the inherent biochemical versatility of members of the ALDH enzyme superfamily in *P. syringae*.

*These authors contributed equally to this work.

Introduction

Aldehyde dehydrogenases (ALDHs[1]) are NAD(P)(H)-dependent enzymes found in prokaryotes and eukaryotes that convert a range of aldehyde substrates into their corresponding carboxylic acids [1–4]. Typically, ALDHs are considered housekeeping enzymes with their biochemical function associated with the general detoxification of reactive aldehydes generated through cellular metabolism [5,6]. The inherent biochemical versatility of the ALDHs allow different members of this enzyme superfamily to contribute
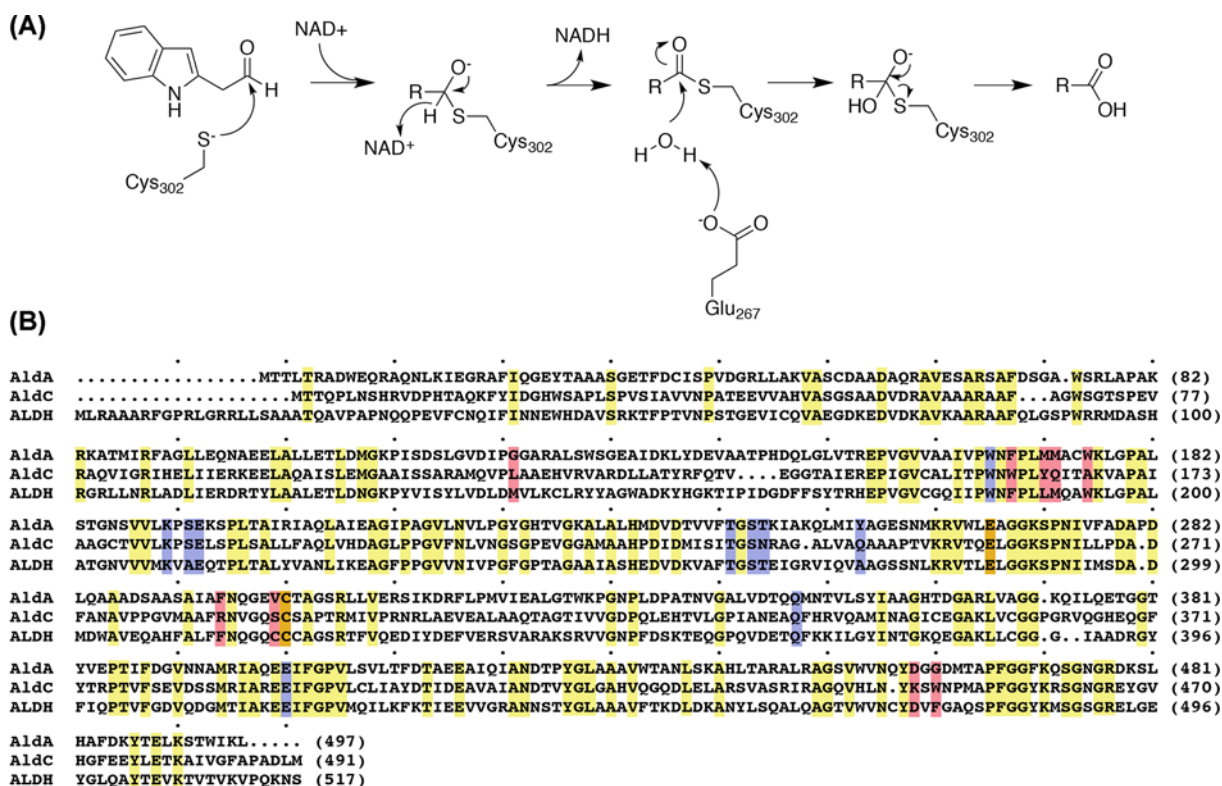
Figure 1. AldA indole-3-acetaldehyde dehydrogenase reaction and sequence comparison

(**A**) Proposed mechanism for IAA synthesis catalyzed by AldA. The general reaction for conversion of an aldehyde into a carboxylic acid applies to other substrates of AldA, including octanal. (**B**) Sequence comparison of AldA (AAO53646.1) and AldC (AAO57114.1) from *P. syringae* strain *Pto*DC3000 and human ALDH (P05091.2). Residues corresponding to the catalytic, NAD(H) binding, and substrate binding sites of AldA are highlighted in orange, blue, and red, respectively. Invariant residues are highlighted in yellow.

to a range of metabolic processes beyond detoxification of aldehydes, including ethanol metabolism [7–9], polyamine metabolism [10], plant cell wall ester synthesis [11,12], metabolism of compounds linked to responses to cellular stresses [13,14], and xenobiotic metabolism [15–17]. Recent studies also suggest that in the plant pathogenic microbe *Pseudomonas syringae* strain *Pto*DC3000 (*Pto*DC3000) an ALDH family member synthesizes indole-3-acetic acid (IAA), the primary auxin hormone of plants, to manipulate host plant auxin responses to promote pathogenicity [18].

P. syringae produces a variety of virulence factors, including phytohormones or chemical mimics of hormones, to manipulate hormone signaling in its host plants as a means of suppressing disease responses and promoting infection [19–21]. Production of the auxin hormone IAA by *P. syringae* and other plant-associated microbial pathogens is implicated in pathogen virulence [18,22–26]. The plant pathogenic *Pto*DC3000 strain synthesizes IAA through the activity of an indole-3-acetaldehyde dehydrogenase [18]. Earlier work by Xie et al. [27] reported a mutation in *Azospirilum brasilense* that decreased IAA levels in the microbe and was shown to affect a gene encoding an ALDH. Subsequent examination of *Pto*DC3000 identified three ALDHs—AldA (UniProt: PSPTO_0092), AldB (UniProt: PSPTO_2673), and AldC (UniProt: PSPTO_3644)—that are related by 30–40% amino acid identity [18].

AldA from *Pto*DC3000 was determined to function as an indole-3-acetaldehyde dehydrogenase through comprehensive metabolic, biochemical, structural, and in planta analyses [18]. AldA was shown to be essential for synthesis of the phytohormone IAA and as a factor contributing to the virulence of *Pto*DC3000 [18]. Although not as metabolically or kinetically efficient as AldA for auxin synthesis, AldB may also contribute to IAA production in this strain [18].

Biochemical and structural studies of AldC from *Pto*DC3000 indicate that this enzyme functions as a long-chain aliphatic ALDH and does not contribute to pathogenicity [18,28]. Structural studies of AldA and AldC from *Pto*DC3000 indicate that both proteins are members of the larger ALDH enzyme superfamily and share a common chemical reaction mechanism, as well as conserved catalytic residues and NAD(H) binding sites with other such enzymes, including human ALDH, but with differences in the substrate binding site (Figure 1) [18,28].

PORTLAND
PRESS

To better understand the biochemical function of the AldA indole-3-acetaldehyde dehydrogenase from *Pto*DC3000, we examined the contribution of putative active site residues (Cys[302] and Glu[267]) on the reaction catalyzed by AldA using a series of site-directed mutants. The 2.85 Å resolution X-ray crystal structure of the catalytically inactive AldA C302A mutant in complex with IAA and NAD[+] reveals that the nicotinamide cofactor adopts a different conformation from that in the previously reported wild-type AldA structure [18]. These structures suggest that AldA undergoes isomerization of the cofactor, as reported for human ALDH [29–30]. Site-directed mutagenesis of the substrate binding site identified Phe[169] as a determinant of substrate selectivity with the AldA F169W mutant shifting substrate preference toward that of AldC, which prefers the 8-carbon aliphatic aldehyde octanal. These are the first biochemical studies examining the structure–function relationship in the AldA indole-3-acetaldehyde dehydrogenase from *Pto*DC3000 and highlight a conserved reaction chemistry and versatility of the active site for accepting a range of substrates.

Materials and methods
Cloning and site-directed mutagenesis
The pET28a-AldA construct was previously described [18]. Site-directed mutants of AldA (catalytic residues: C302A, E267Q, E267A; substrate-binding site: G123L, F169A, F169W, M172A, M173A, W176A, F296A, V301A, D459K, G461W) were generated using the QuikChange PCR method (Agilent Technologies) with the pET28a-AldA construct as template. For expression of His$_6$-tagged AldA proteins expression, each pET28a-AldA construct was transformed into *Escherichia coli* BL21 (DE3) (EMD Millipore).

Protein expression and purification
Transformed *E. coli* BL21 (DE3) cells containing either wild-type or mutant AldA construct were grown at 37°C in Terrific broth with 50 μg.ml^{-1} kanamycin until A$_{600nm}$ ~0.8. Protein expression was induced by addition of 1 mM isopropyl-1-thio-β-D-galactopyranoside to the culture and the cells were then grown at 16°C overnight. Cells were harvested by centrifugation (5000×g for 30 min) and the cell pellet resuspended in 50 mM Tris, pH 8.0, 500 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, and 1% (v/v) Tween-20. The resuspended cells were lysed by sonication with cell debris removed by centrifugation (12000×g for 45 min). The resulting supernatant was loaded on to a 2–4 ml Ni^{2+}-nitriloacetic acid (NTA) column (Qiagen). The column was washed with three to five column volumes of 50 mM Tris, pH 8.0, 500 mM NaCl, 25 mM imidazole, and 10% (v/v) glycerol to remove unbound proteins. The His$_6$—AldA fusion protein was eluted from the Ni^{2+}-NTA column using 50 mM Tris, pH 8.0, 500 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, and 250 mM imidazole. Eluted His$_6$-tagged AldA protein (either wild-type or mutant) was further purified using 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 protein were pooled and concentrated to 8–10 mg.ml^{-1}. Protein concentrations were determined using the Bradford method with bovine serum albumin as a standard.

Enzyme assays and steady-state kinetic analysis
Enzymatic activity of wild-type and mutant AldA proteins was measured by monitoring NADH formation (ε340 nm = 6220 M^{-1}.cm^{-1}; 100 μl volume) at A$_{340nm}$ using an EPOCH2 microplate spectrophotometer (BioTek), as reported previously [18]. Experiments were performed at 25°C in a standard assay mix of 100 mM Tris/HCl (pH 8.0) and 100 mM KCl. Initial activity assays used standard assay conditions with fixed concentrations of NAD$^+$ (2 mM) and either indole-3- acetaldehyde (10 mM) or octanal (10 mM). Steady-state kinetic parameters of wild-type and mutant AldA proteins were determined at 25°C in a standard assay mix with fixed NAD$^+$ (2.0 mM) and either varied indole-3-acetaldehyde (0–2.5 mM) or varied octanal (0–10 mM). Protein amounts ranged from 0.1 μg for wild-type AldA to 100 μg for less active point mutants. The resulting initial velocity data were fit to the Michaelis–Menten equation, $v = (k_{cat} [S])/(K_m + [S])$, using SigmaPlot.

Protein crystallography
Protein crystals of the AldA C302A mutant were grown by the hanging drop vapor diffusion method at 4°C. Crystals of the AldA C302A mutant (9.8 mg.ml^{-1}) in complex with IAA and NAD$^+$ formed grew in drops of a 1:1 mixture of proteins and crystallization buffer, which was 24% (w/v) PEG-1000, 100 mM Tris/HCl (pH 7.0), 2 mM IAA, and 5 mM NAD$^+$. Crystals were stabilized in mother liquor with 30% (v/v) glycerol added as a cryoprotectant before flash-freezing in liquid nitrogen for data collection at 100 K. Diffraction data were collected at beamline 19-ID of the Structural Biology Center, Advanced Photon Source, Argonne National Lab. HKL3000 was used to index, integrate, and scale the collected X-ray data [31]. Molecular replacement was used to solve the X-ray crystal structure of

Table 1 Summary of crystallographic data collection and refinement statistics

Crystal	AldA (C302A)•IAA•NAD$^+$
Space group	C2
Cell dimensions	$a = 333.8$ Å, $b = 161.1$ Å, $c = 85.0$ Å
Data collection	
Wavelength	0.979 Å
Resolution range (highest shell)	47.9–2.85 Å (2.92–2.85 Å)
Reflections (total/unique)	555535/103193
Completeness (highest shell)	98.3% (99.8%)
$<I/\sigma>$ (highest shell)	13.9 (5.4)
R_{sym} (highest shell)	9.2% (24.1%)
Refinement	
R_{cryst}/R_{free}	21.6%/25.4%
Number of protein atoms	29536
Number of ligand atoms	456
R.m.s. deviation, bond lengths	0.010 Å
R.m.s. deviation, bond angles	1.18°
Avg. B-factor: protein, ligand	35.9, 69.4 Å^2
Stereochemistry: favored, allowed, outliers	95.8, 4.2, 0%

AldA using PHASER [32] with the three-dimensional structure of wild- type AldA PDB: 5IUW) [18]. COOT [33] and PHENIX [34] were used for iterative rounds of model building and refinement, respectively. Data collection and refinement statistics are summarized in Table 1. The final model consisted of eight protein monomers in the asymmetric unit, each with IAA and NAD$^+$ bound. No waters were added during refinement due to the resolution of the structure. Atomic coordinates and structure factors for the AldA (C302A)•IAA•NAD$^+$ complex (PDB: 7JSO) were deposited in the RCSB Protein Data Bank (www.rcsb.org).

Results and discussion
Mutagenesis of the catalytic cysteine and glutamate residues in AldA

Initial biochemical analysis of AldA showed that this ALDH converts indole-3-acetaldehyde into IAA and shares a conserved three-dimensional structure and active site features with other members of the ALDH enzyme superfamily [18]. As with other ALDHs [4,8,35–37], AldA retains the invariant active site cysteine (Cys^{302}) required for formation of a thioacyl-enzyme intermediate in the reaction sequence and a conserved glutamate (Glu^{267}) that serves to activate a water molecule for hydrolysis of the intermediate (Figure 1). To probe the role of the corresponding catalytic residues in AldA, the C302A, E267Q, and E267A site-directed mutants were generated. Each mutant was expressed in *E. coli* as a His$_6$-tagged protein with yields and a tetrameric oligomerization similar to that of wild-type AldA [18] (Supplementary Figure S1). As reported previously [18], wild-type AldA was active with indole-3-acetaldehyde as a substrate (Table 2). Enzyme assays using the AldA C302A, E267Q, and E267A mutants showed no detectable activity above background rates in assays using up to 1000-fold higher amounts of protein compared with wild-type AldA. This result is consistent with the canonical ALDH reaction mechanism in which a cysteine (Cys^{302} in AldA) reacts with the aldehyde substrate to form an thioacyl-enzyme intermediate, which reacts with NAD$^+$ for hydride transfer, followed by hydrolysis of the intermediate via a water molecule activated by a glutamate (Glu^{267} in AldA) and subsequent release of the carboxylic acid product [8,19,35–37].

## Structure of AldA C302A mutant in complex with IAA and NAD$^+$ and cofactor isomerization

To further examine the reaction mechanism of AldA, the AldA C302A mutant was crystallized in the presence of NAD$^+$ and IAA. The 2.85 Å resolution X-ray crystal structure of the AldA C302A mutant in complex with these ligands was solved by molecular replacement (Table 1). The overall three-dimensional structure of the AldA C302A mutant is a tetramer (Figure 2A) with each monomer consisting of an N-terminal NAD(H) binding domain defined by a canonical Rossmann-fold, an interdomain linker, a β-strand oligomerization region, and a C-terminal mixed $^\alpha$/β-domain that includes the catalytic cysteine and the aldehyde substrate binding site (Figure 2B). The AldA C302A mutant structure is similar to that of the previously reported wild-type protein [18] with a 0.24 Å^2 r.m.s.d. for 497

PORTLAND PRESS

Table 2 Steady-state kinetic parameters of wild-type and mutant AldA proteins

Protein	Indole-3-acetaldehyde			Octanal		
	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1}.s^{-1})	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1}.s^{-1})
WT	240 $\pm$ 9	0.18 $\pm$ 0.02	22,040	96 $\pm$ 6	2.7 $\pm$ 0.6	593
G123L	202 $\pm$ 7	0.26 $\pm$ 0.04	12,950	71 $\pm$ 4	2.5 $\pm$ 0.5	473
F169A	88 $\pm$ 12	4.7 $\pm$ 0.6	312	50 $\pm$ 4	>10*	83
F169W	41 $\pm$ 4	1.6 $\pm$ 0.4	427	108 $\pm$ 13	0.69 $\pm$ 0.10	2609
M172A	220 $\pm$ 23	0.92 $\pm$ 0.11	3986	92 $\pm$ 4	2.7 $\pm$ 2.5	567
M173A	184 $\pm$ 23	1.5 $\pm$ 0.4	2044	63 $\pm$ 9	>10*	105
W176A	10 $\pm$ 2	8.8 $\pm$ 2.5	19	-	-	-
F296A	18 $\pm$ 1	7.0 $\pm$ 2.1	43	8	>10*	13
V301A	81 $\pm$ 10	7.2 $\pm$ 0.9	190	32 $\pm$ 9	>10*	54
D459K	72 $\pm$ 14	0.86 $\pm$ 0.23	1395	–	–	–
G461W	160 $\pm$ 14	0.29 $\pm$ 0.05	9195	98 $\pm$ 5	4.1 $\pm$ 0.8	398

The protein and location of the mutation are indicated.
Enzyme assays were performed as described in the 'Materials and methods' section.
Average values $\pm$ S.D. (*n*=3) are shown.
*Estimated K_m greater than 10 mM solubility of substrate tested.
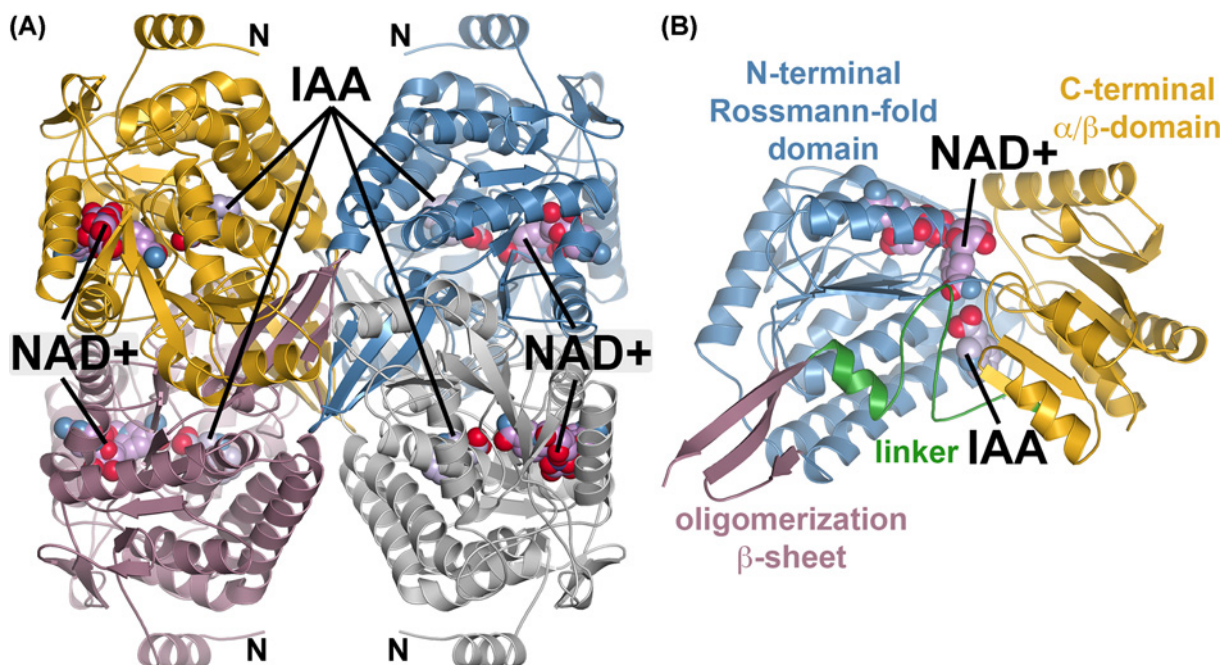

Figure 2. Overall three-dimensional structure of the AldA(C302A)•NAD$^+$•IAA complex
(**A**) Tetrameric assembly of the AldA C302 mutant. Each monomer is colored individually with the locations of bound NAD$^+$ and IAA indicated. The N-terminus of each monomer is also noted. (**B**) Domain organization of an AldA monomer. The view is slightly rotated from that in (A). The N-terminal Rossmann-fold (blue), C-terminal $^\alpha$/β domain (gold), oligomerization β-sheet (rose), and domain linker (green) are indicated and labeled with ligand positions indicated.

C$_\alpha$-atoms in the monomer. NAD$^+$ was observed bound in the N-terminal NAD(H) binding site and IAA in the C-terminal substrate binding site.

Although the overall three-dimensional structures of the wild-type and C302A mutant AldA proteins are similar, there is a difference in the observed conformations of NAD$^+$ in each structure (Figure 3). As reported previously, residues in the NAD(H) binding site of AldA and other ALDHs are highly conserved (Figure 1B). In the AldA C302A mutant structure, the adenine half of the NAD$^+$ molecule is bound in the same orientation as previously reported for the wild-type enzyme [18]. Likewise, the same interactions between the adenine-ribose-pyrophosphate moieties of
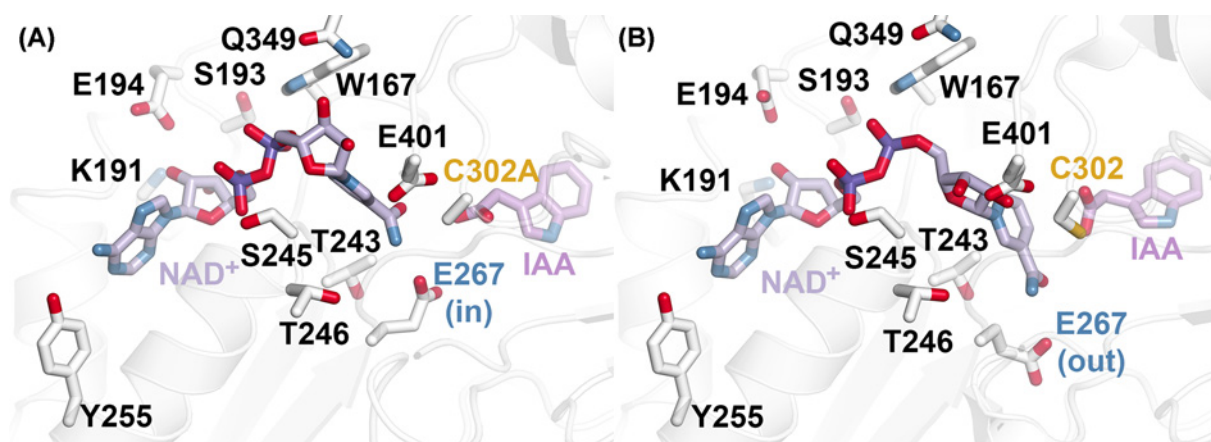
Figure 3. NAD$^+$ isomerization in AldA

(**A**) NAD$^+$ binding in the AldA C302A X-ray crystal structure. The nicotinamide half of the ligand adopts a conformation that positions the nicotinamde ring away from the catalytic site (i.e., C302A in gold), which is ideal for the hydrolysis step of the reaction. The side-chain of Glu^{267} (blue; in) is positioned into the active site. (**B**) NAD$^+$ binding in the AldA wild-type X-ray crystal structure. As previously reported [18], the nicotinamide half of the ligand adopts a conformation that positions the nicotinamde ring into the catalytic site (i.e., C302 in gold), which is ideal for the hydride transfer step of the reaction. The side-chain of Glu^{267} (blue; out) is positioned away from the active site.

NAD$^+$ and Tyr^{255}, Lys^{191}, Glu^{194}, Ser^{193}, Ser^{245}, and Trp^{167} are observed in the mutant AldA structure (Figure 3). The nicotinamide half of the coenzyme is bound with the ribose oriented toward Gln^{349} and the nicotinamide ring pulled away from the catalytic site (defined by the C302A mutation) by contraction of the coenzyme structure (Figure 3A). This orientation of NAD$^+$ retains a hydrogen bond between the ligand amide group and the carboxylate side-chain of Glu^{267}, which is positioned into the AldA active site. In comparison, in the earlier X-ray crystal structure of AldA [18] NAD$^+$ adopted an extended conformation with the nicotinamide half of the cofactor placed deeper into the active site and in proximity to Cys^{302} (Figure 3B). This allows Glu^{401} to form a bi-dendate interaction with the hydroxyl groups of the nicotinamide-ribose group. The nicotinamide ring is positioned in proximity to the catalytic cysteine with the amide moiety of NAD$^+$ still hydrogen bonding with the carboxylate side-chain of Glu^{267}, which shifts outward from the catalytic site (Figure 3B).

Site-directed mutagenesis of the conserved catalytic residues of AldA and the three-dimensional structures of this enzyme imply a shared reaction mechanism with other members of the ALDH superfamily (Figure 1A), including coenzyme isomerization that allows for hydride transfer (Figure 3B) and hydrolysis of the thioacyl-enzyme intermediate (Figure 3A) during the reaction mechanism [29,35]. As for other ALDHs, we suggest that AldA binds indole-3-acetaldehyde and NAD$^+$ in its active site and undergoes formation of a thioacyl-enzyme intermediate at Cys^{302} and subsequent hydride transfer (Figure 1A) with the cofactor adopting a compact conformation similar to that observed in the AldA C302A mutant structure (Figure 3). Isomerization of NAD$^+$ and movement of Glu^{267} then allows for hydrolysis by an activated water molecule and release of the final IAA product.

Site-directed mutagenesis of the AldA substrate-binding site and identification of a determinant of substrate selectivity

The three-dimensional structures of both wild-type AldA [18] and the C302A mutant (Figure 4A) provide similar information on how IAA binds in the active site. The IAA carboxylate group is oriented into the catalytic site (i.e., ∼3Å from side-chain of C302A in AldA) with the indole ring positioned by van der Waals interactions in a largely apolar binding site, which includes Gly^{123}, Phe^{169}, Met^{172}, Met^{173}, Trp^{176}, Phe^{296}, Val^{301}, Thr^{303}, Asp^{459}, and Gly^{461} (Figure 4A). To probe the contributions of various residues in the AldA binding site, a series of site-directed point mutants (G123L, F169A, F169W, M172A, M173A, W176A, F296A, V301A, D459K, and G461W) were generated. The G123L, F169W, D459K, and G461W mutants substitute the AldA residue with the corresponding amino acid found in AldC (Figure 1B). Each AldA mutant was expressed as a His$_6$-tagged fusion protein in *E. coli* and isolated by affinity chromatography with yields comparable with wild-type.
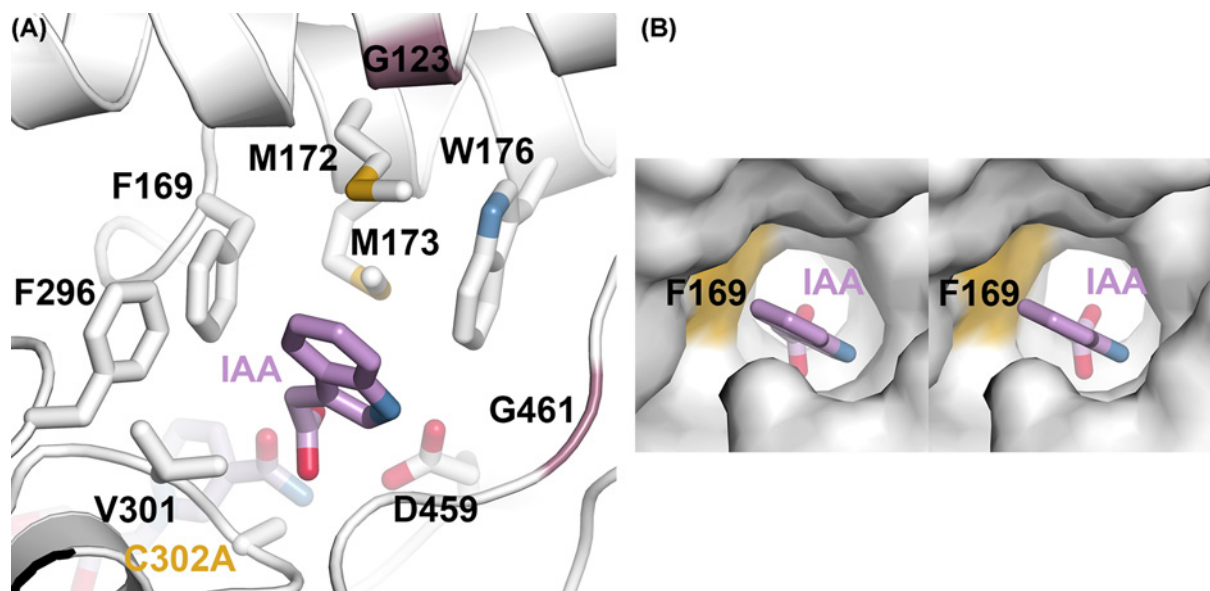
Figure 4. AldA C302A mutant substrate binding site

(**A**) IAA binding in the AldA C302A mutant X-ray crystal structure. Residues in the substrate binding site are shown. Two glycine residues in the site are represented by the rose colored secondary structure. (**B**) Surface stereo-view of the AldA substrate binding site. The surface corresponding to Phe[169] is highlighted in gold.

Using indole-3-acetaldehyde as a substrate (Table 2), AldA variants with mutations in residues closer to the solvent-exposed entrance of the IAA binding site—G123L and G461W—displayed approximately two-fold decreases in catalytic efficiency (i.e., k_{cat}/K_m). Mutations of Met[172] (M172A), Met[173] (M173A), and Asp[459] (D459K) resulted in 5- to 15-fold reductions in k_{cat}/K_m compared with wild-type AldA. The most detrimental substitutions were those of residues in the 'aromatic box'—Phe[169], Trp[176], Phe[296], and Val[301]. The F169A, F169W, W176A, F296A, and V301A mutants generally exhibited decreased turnover rates and increased K_m values with roughly 50- to 1150-fold decreases in catalytic efficiency. These results suggest that alterations in the apolar surface of the AldA substrate site likely alter binding of indole-3-acetaldehyde for catalysis.

Recent work identified the AldC enzyme from *P. syringae* strain *Pto*DC3000 as a long-chain aliphatic ALDH [28]. To compare the activity of AldA with AldC, wild-type AldA was assayed using octanal as a substrate; this is the preferred substrate of AldC [28]. Although AldA displayed activity with octanal (Table 2 and Supplementary Figure S2A,B), it was 37-fold less efficient as a substrate compared with indole-3-acetaldehyde. We also used octanal as a substrate for the AldA point mutants described above (Table 2). Two mutants—W176A and D459K—had no detectable activity with this substrate. The F169A, M173A, F296A, and V301A mutants exhibited activity with octanal, but no substrate saturation. This indicates that the potential K_m values of these mutants are greater than the 10 mM solubility limit of octanal. The AldA G123L, M172A, and G461W substitutions had only modest effect on the catalytic efficiency with octanal compared with wild-type protein. Interestingly, the F169W mutant accepted octanal (k_{cat}/K_m = 2610 M^{-1}.s^{-1}) as the preferred substrate compared with indole-3-acetaldehyde (k_{cat}/K_m = 427 M^{-1}.s^{-1}) (Table 2 and Supplementary Figure S2C,D). This mutation altered substrate selectivity of AldA toward that of AldC by six-fold.

The three-dimensional structure of AldA suggests a possible basis for the change in substrate selectivity of the AldA F169W mutant (Figure 4B). In AldA, Phe[169] is positioned deep in the 'aromatic box' region of the substrate binding site. As described for other ALDHs, residues in this region of the active site form an apolar tunnel that influences substrate selectivity [38–40].

Substitution of Phe[169] with a larger tryptophan residue likely constricts the entrance to the catalytic center of AldA. The tryptophan side-chain of the F169W mutant may sterically interfere with the indole ring of IAA; whereas, the aliphatic chain of octanal better fits the substrate binding site of the mutant. This would account for the change in selectivity for octanal as a substrate versus indole-3-acetaldehyde in the AldA F169W mutant. Interestingly, the corresponding residue at this position in AldC from *P. syringae*, which prefers octanal versus other aliphatic aldehydes as a substrate, is a tryptophan [28]. Comparison of the k_{cat}/K_m for the AldA F169W mutant (2610 M^{-1}s^{-1}) with that

of AldC (924 $M^{-1}s^{-1}$; [28]) suggest that the identity of the residue at this position of the substrate binding site of these enzymes is critical for substrate preference.

Summary

As a widespread enzyme family, ALDHs catalyze conversion of aldehydes into carboxylates using a shared NAD(H)-dependent mechanism [1–4]. The versatility of these enzymes allows them to perform a range of biological roles. For example, in the plant pathogen *P. syringae* uses AldA as an indole-3-acetaldehyde dehydrogenase that contributes to pathogen virulence [18]. Here we further examined the biochemical function of AldA. Site-directed mutagenesis of AldA confirms that the enzyme shares the same catalytic residues (Cys^{302} and Glu^{267}) as other ALDHs. As described for other ALDHs, the cysteine reacts with the aldehyde to form a thioacyl-enzyme intermediate, which reacts with NAD^+ for hydride transfer, followed by hydrolysis of the intermediate via a water molecule activated by a glutamate, and release of the carboxylate (Figure 1) [8,19,35–37].

Structural analysis of the AldA C302A mutant and comparison with the wild-type AldA (Figure 3) reveals that the NAD^+ undergoes a conformational isomerization, as reported for other ALDHs [29,35], that allows for each half reaction to occur. The other common feature of the ALDH family is the inherent versatility of their substrate binding sites to accommodate a variety of molecules, as these enzymes typically display broad substrate profiles [2–7,38–40]. For example, the 'aromatic box' formed by apolar residues in the substrate binding site helps define which aldehydes are recognized by the ALDH [38–40]. Although mutations in the AldA substrate binding site reduced activity with indole-3-acetaldehyde, as well as octanal (Table 2), substitution of Phe^{169} with a tryptophan (the residue at this position found in AldC) enhanced activity with the aliphatic aldehyde. This suggests that key changes in the site are important for the evolution of the diverse biochemical functions in members of the ALDH enzyme superfamily.

Data Availability

The data and material presented in this manuscript are available from the corresponding author on reasonable request. Atomic coordinates and structure factors for the AldA (C302A)•IAA•NAD^+ complex (PDB: 7JSO) were deposited in the RCSB Protein Data Bank (www.rcsb.org).

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

Author Contribution

K.Z., J.S.L., R.L., Z.T.C., T.J.D., E.S.D.T., C.T.E., I.K.E, A.R.G., L.A.G., E.M.H., K.A.H., K.H., A.K., K.C.K., A.Z.L., J.Y.L., A.Y.L., O.W.P., E.P.R., S.S., M.L.S., P.L.W., G.J.W., M.C.X., R.S.Y., B.C., W.C., and C.K.H. generated site-directed mutants, purified protein, and performed enzyme assays. R.L. and C.K.H. purified protein, collected X-ray diffraction data, and refined the X-ray crystal structures. B.C., W.C., C.K.H., S.A.M., B.N.K., S.G.L., and J.M.J. contributed to the design of the experiments and data analyses. J.M.J wrote the manuscript with all authors providing feedback.

Abbreviations

ALDH, aldehyde dehydrogenase; IAA, indole-3-acetic acid; NTA, nitriloacetic acid; *Pto*DC3000, *Pseudomonas syringae* strain *Pto*DC3000.

References

1   Yoshida, A., Rzhetsky, A., Hsu, L.C. and Chang, C. (1998) Human aldehyde dehydrogenase gene family. *Eur. J. Biochem.* **251**, 549–557, https://doi.org/10.1046/j.1432-1327.1998.2510549.x

2   Brocker, C., Vasiliou, M., Carpenter, S., Carpenter, C., Zhang, Y., Wang, X. et al. (2013) Aldehyde dehydrogenase (ALDH) superfamily in plants: gene nomenclature and comparative genomics. *Planta* **237**, 189–210, https://doi.org/10.1007/s00425-012-1749-0

3 Sophos, N.A. and Vasiliou, V. (2003) Aldehyde dehydrogenase gene superfamily: the 2002 update. *Chem. Biol. Interact.* **143–144**, 5–22, https://doi.org/10.1016/S0009-2797(02)00163-1

4 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, https://doi.org/10.1016/j.cbi.2019.03.006

5 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, https://doi.org/10.1080/10408440591002183

6 Vasiliou, V., Pappa, A. and Estey, T. (2004) Role of human aldehyde dehydrogenases in endobiotic and xenobiotic metabolism. *Drug Metab. Rev.* **36**, 279–299, https://doi.org/10.1081/DMR-120034001

7 Klyosov, A.A. (1996) Kinetics and specificity of human liver aldehyde dehydrogenases toward aliphatic, aromatic, and fused polycyclic aldehydes. *Biochemistry* **35**, 4457–4466, https://doi.org/10.1021/bi9521102

8 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, https://doi.org/10.1016/S0969-2126(97)00224-4

9 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, https://doi.org/10.1110/ps.8.12.2784

10 Brocker, C., Lassen, N., Estey, T., Pappa, A., Cantore, M., Orlova, V. et al. (2010) Aldehyde dehydrogenase 7A1 (ALDH7A1) is a novel enzyme involved in cellular defense against hyperosmotic stress. *J. Biol. Chem.* **285**, 18452–18463, https://doi.org/10.1074/jbc.M109.077925

11 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, https://doi.org/10.1105/tpc.017509

12 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, https://doi.org/10.1093/jxb/err045

13 Rodrigues, S.M., Kotchoni, S.O., Kuhns, C., Ditzer, 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, https://doi.org/10.1111/j.1365-3040.2005.01458.x

14 Rodrigues, S.M., Andrade, M.O., Gomes, A.P., Damatta, F.M., Baracat-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, https://doi.org/10.1093/jxb/erj132

15 Coitinho, J.B., Pereira, M.S., Costa, D.M., Guimarães, S.L., Araújo, S.S., Hengge, A.C. et al. (2016) Structural and kinetic properties of the aldehyde dehydrogenase NahF, a broad substrate specificity enzyme for aldehyde oxidation. *Biochemistry* **55**, 5453–5463, https://doi.org/10.1021/acs.biochem.6b00614

16 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, https://doi.org/10.1016/j.abb.2017.01.011

17 Zahniser, M.P.D., Prasad, S., Kneen, M.M., Kreinbring, C.A., Petsko, G.A., Ringe, D. et al. (2017) Structure and mechanism of benzaldehyde dehydrogenase from *Pseudomonas putid*a ATCC 12633, a member of the class 3 aldehyde dehydrogenase superfamily. *Protein Eng. Des. Sel.* **30**, 271–278, https://doi.org/10.1093/protein/gzx015

18 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 syringa*e strain DC 3000. *PLoS Pathog.* **14**, e1006811, https://doi.org/10.1371/journal.ppat.1006811

19 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, https://doi.org/10.1146/annurev-phyto-082712-102321

20 Jones, J.D.G. and Dangl, J.L. (2006) The plant immune system. *Nature* **444**, 323–329, https://doi.org/10.1038/nature05286

21 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, https://doi.org/10.1016/j.cell.2006.02.008

22 Djami-Tchatchou, A.T., Harrison, G.A., Harper, C.P., Wang, R., Prigge, M., Estelle, M. et al. (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, https://doi.org/10.1094/MPMI-02-20-0047-R

23 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, https://doi.org/10.1093/jxb/erx447

24 Aragon, I.M., Perez-Martinez, I., Moreno-Perez, 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, https://doi.org/10.1111/1574-6968.12413

25 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, https://doi.org/10.1111/tpj.12157

26 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. gypsophilae. *Mol. Plant Microbe Interact.* **11**, 634–642, https://doi.org/10.1094/MPMI.1998.11.7.634

27 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, https://doi.org/10.1016/j.femsle.2005.05.020

28 Lee, S.G., Harline, K., Abar, O., Akadri, S.O., Bastian, A.G., Chen, H.Y. et al. (2020) Biochemical and structural analysis of a long-chain aliphatic aldehyde dehydrogenase (AldC) from the plant pathogen *Pseudomonas syringae* DC3000. *J. Biol. Chem.* **295**, 13914–13926, https://doi.org/10.1074/jbc.RA120.014747

29 Perez-Miller, S.J. and Hurley, T.D. (2003) Coenzyme isomerization is integral to catalysis in aldehyde dehydrogenase. *Biochemistry* **42**, 7100–7109, https://doi.org/10.1021/bi034182w

30 Talfournier, F., Pailot, A., Stinès-Chaumeil, C. and Branlant, G. (2009) Stabilization and conformational isomerization of the cofactor during the catalysis in hydrolytic ALDHs. *Chem. Biol. Interact.* **178**, 79–83

31 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, https://doi.org/10.1107/S0907444906019949

32 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. Cryst.* **40**, 658–674, https://doi.org/10.1107/S0021889807021206

33 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132, https://doi.org/10.1107/S0907444904019158

34 Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I. et al. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877, https://doi.org/10.1107/S2059798319011471

35 Marchal, S., Rahuel-Clermont, S. and Branlant, G. (2000) Role of glutamate-268 in the catalytic mechanism of nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase from *Streptococcus mutans*. *Biochemistry* **39**, 3327–3335

36 D'Ambrosio, K., Pailot, A., Talfournier, F., Didierjean, C., Benedetti, E., Aubry, A. et al. (2006) The first crystal structure of a thioacylenzyme intermediate in the ALDH family: new coenzyme conformation and relevance to catalysis. *Biochemistry* **45**, 2978–2986, https://doi.org/10.1021/bi0515117

37 Munoz-Clares, R.A., Gonzalez-Segura, L., Riveros-Rosas, H. and Julian-Sanchez, A. (2015) Amino acid residues that affect the basicity of the catalytic glutamate of the hydrolytic aldehyde dehydrogenases. *Chem. Biol. Interact.* **234**, 45–58, https://doi.org/10.1016/j.cbi.2015.01.019

38 Riveros-Rosas, H., Gonzalez-Segura, L., Julian-Sanchez, A., Diaz-Sanchez, A.G. and Munoz-Clares, R.A. (2013) Structural determinants of substrate specificity in aldehyde dehydrogenases. *Chem. Biol. Interact.* **202**, 51–61, https://doi.org/10.1016/j.cbi.2012.11.015

39 Luo, M. and Tanner, J.J. (2015) Structural basis of substrate recognition by aldehyde dehydrogenase 7A1. *Biochemistry* **54**, 5513–5522, https://doi.org/10.1021/acs.biochem.5b00754

40 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, https://doi.org/10.1016/j.abb.2013.07.024