Structural and kinetic features of aldehyde dehydrogenase 1A (ALDH1A) subfamily members, cancer stem cell markers active in retinoic acid biosynthesis

Raquel Pequerul[1], Javier Vera[1], Joan Giménez-Dejoz[1], Isidro Crespo[1], Joan Coines[2], Sergio Porté[1], Carme Rovira[2], Xavier Parés[1], Jaume Farrés[1]*

From the [1]Department of Biochemistry and Molecular Biology, Faculty of Biosciences, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Barcelona, Spain; [2]Department of Inorganic and Organic Chemistry, Faculty of Physics and Chemistry, Universitat de Barcelona E-08028 Barcelona, Spain

*To whom correspondence should be addressed: Jaume Farrés: Department of Biochemistry and Molecular Biology, Faculty of Biosciences, E-08193 Bellaterra (Barcelona) Spain; jaume.farres@uab.cat; Tel: +34-93-581 2557; Fax: +34-93-581 1264

ABSTRACT

Aldehyde dehydrogenases catalyze the $NAD(P)^+$-dependent oxidation of aldehydes to their corresponding carboxylic acids. The three-dimensional structures of the human ALDH1A enzymes were recently obtained, while a complete kinetic characterization of them, under the same experimental conditions, is lacking. We show that the three enzymes, ALDH1A1, ALDH1A2 and ALDH1A3, have similar topologies, although with decreasing volumes in their substrate-binding pockets. The activity with aliphatic and retinoid aldehydes was characterized side-by-side, using an improved HPLC-based method for retinaldehyde. Hexanal was the most efficient substrate. ALDH1A1 displayed lower K_m values with hexanal, *trans*-2-hexenal and citral, compared to ALDH1A2 and ALDH1A3. ALDH1A2 was the best enzyme for the lipid peroxidation product, 4-hydroxy-2-nonenal, in terms of k_{cat}/K_m. The catalytic efficiency towards all-*trans* and 9-*cis*-retinaldehyde was in general lower than for alkanals and alkenals. ALDH1A2 and ALDH1A3 showed higher catalytic efficiency for all-*trans*-retinaldehyde. The lower specificity of ALDH1A3 for 9-*cis*-retinaldehyde against the all-*trans*- isomer might be related to the smaller volume of its substrate-binding pocket. Magnesium inhibited ALDH1A1 and ALDH1A2, while it activated ALDH1A3, which is consistent with cofactor dissociation being the rate-limiting step for ALDH1A1 and ALDH1A2, and deacylation for ALDH1A3, with hexanal as a substrate. We mutated both ALDH1A1 (L114P) and ALDH1A2 (N475G, A476V, L477V, N478S) to mimic their counterpart substrate-binding pockets. ALDH1A1 specificity for citral was traced to residue 114 and to residues 458 to 461. Regarding retinaldehyde, the mutants did not show significant differences with their respective wild-type forms, suggesting that the mutated residues are not critical for retinoid specificity.

Keywords: aldehyde dehydrogenase, all-*trans*-retinaldehyde, magnesium, substrate-binding pocket, retinoic acid.

ABBREVIATIONS: ALDH, Aldehyde dehydrogenase; BSA, Bovine serum albumin; DTT, Dithiothreitol; EDTA, Ethylenediaminetetraacetic acid; FPLC, Fast protein liquid chromatography; HPLC, High-performance liquid chromatography; IPTG, Isopropyl β-D-1-thiogalactopyranoside; PMSF, Phenylmetane sulfonyl fluoride; RA, Retinoic acid; RAL, Retinaldehyde; RALDH, Retinaldehyde dehydrogenase; SDS-PAGE, Sodium dodecyl sulfate-polyacrylamide gel electrophoresis; Triton X-100, 4-(1,1,3,3-Tetramethylbutyl) phenyl-polyethylene glycol.

Introduction

The aldehyde dehydrogenase (ALDH) gene superfamily (E.C. 1.2.1.3) encodes a cluster of evolutionarily related $NAD(P)^+$-dependent enzymes catalyzing the irreversible oxidation of a wide spectrum of aldehyde substrates, generated from various endogenous and exogenous precursors, to their corresponding carboxylic acids. Aldehydes are formed during the metabolism of amino acids, biogenic amines, vitamins, steroids and retinoids. Moreover, aldehydes are often generated during the metabolism of xenobiotics such as drugs, food components and environmental agents. Although most ALDHs display broad specificities, oxidizing a variety of both aliphatic and aromatic aldehydes, several forms possess narrower substrate preferences [1]. Aside from their role in aldehyde detoxification, many ALDH show multiple additional catalytic and non-catalytic functions, as the catalysis of ester hydrolysis, and as binding proteins for various endogenous (e.g., hormones) and exogenous (e.g., acetaminophen) compounds [2], and antioxidant roles [3–7].

The human ALDH superfamily consists of 19 putatively functional genes with distinct chromosomal locations. The human ALDH1A subfamily (ALDH1A1, ALDH1A2 and ALDH1A3) are cytosolic tetrameric enzymes with 55-kDa subunits and share 70% amino acid sequence identity. Several crystal structures of human ALDH1A1, ALDH1A2 and ALDH1A3 have been deposited recently (Supplementary material, Table S1) [8–10]. Each ALDH monomer has a similar fold containing three domains: NAD^+-binding, catalytic and oligomerization domains. At the intersection of the substrate- and cofactor-binding pockets sits an strictly conserved cysteine residue (Cys303, ALDH1A1 numbering), which is essential for catalysis, acting as the active-site nucleophile [11]. Phylogenetic analysis shows that the ALDH1A1 and ALDH1A2 forms are more closely related, while ALDH1A3 would be the most divergent. They are highly conserved in mammals, but an ALDH1A1 orthologue is missing in zebrafish [12]. A special interest of these ALDHs is their role in the physiological oxidation of retinaldehyde to retinoic acid (RA).

RA is a potent hormone regulating embryonic development and cell differentiation. RA regulates gene expression and exerts its pleiotropic effects mainly through binding to two classes of nuclear ligand-dependent receptors, RAR and RXR [13]. RA is mostly considered a differentiating and antitumor agent and its production by ALDH can induce differentiation of hematopoietic and neural stem cells. This is the basis behind the successful RA treatment for acute promyelocytic leukemia. Clinical trials are investigating the use of retinoids in the prevention and treatment of several cancer types [14]. RA and its derivatives are also widely used for the treatment of skin related diseases [15].

The three ALDH1A enzymes have been identified in human, mouse and Xenopus, with different physiological functions in embryogenesis [16]. Knockouts of ALDH1A forms severely affect the development, especially the ALDH1A2 and ALDH1A3 mutants, being therefore indispensable for the formation of the precise spatio-temporal concentration gradients of RA in the embryo. ALDH1A enzymes are good markers of cancer stem cells and have been implicated in different types of cancer and in chemoresistance [17].

Although some kinetic constants of human ALDH1A1, ALDH1A2 and ALDH1A3 with retinaldehyde have been reported [9,18,19], a complete side-by-side characterization of the three human ALDH1A enzymes using retinoids and other relevant physiological substrates is lacking. In previous studies, the kinetic properties of human ALDH1A enzymes with retinoids were obtained under various conditions, in different laboratories and in the presence of detergent. Our results, obtained under more physiological conditions, will give a better estimate of the relative contribution of each enzyme in the metabolism of aldehydes, particularly retinaldehyde. In addition, given the fact that several cancer types express one or more of these ALDH1A forms, a full kinetic characterization of each individual form is warranted to undertake pharmacological approaches.

In the present study we carried out the enzymatic characterization of the three human ALDH1A enzymes. Firstly, their three-dimensional structures were compared to recognize the differences that could account for the specific kinetic features. The enzymes were then recombinantly expressed in a heterologous system and purified as soluble, enzymatically active forms. The optimal Mg^{2+} concentration

was determined for each enzymatic form and a complete study of their kinetic parameters, with non-retinoid substrates as well as retinaldehyde isomers, was performed. For the latter substrates, an improved HPLC-based methodology was used for the first time to follow the ALDH activity. The role of several active-site residues was explored by site-directed mutagenesis.

Experimental procedures

Materials

Vectors harboring ALDH cDNAs, pOTB7 (ALDH1A1 and ALDH1A3) and pBluescriptR (ALDH1A2), were purchased from Thermo Scientific. All PCR primers and restriction enzymes were obtained from Invitrogen. The KOD Hot Start polymerase and the DNA ligase kit were purchased from Novagen and Roche, respectively. Expression vector pET-30 Xa/LIC was from Novagen. Host DH5α and BL21 (DE3) pLys cells were obtained from Fermentas and Novagen, respectively. PMSF, retinoids and other substrates were purchased from Sigma-Aldrich. Linear and supercoiled DNA ladders were provided by Thermo Scientific, and Bench™ Protein ladder was purchased from Invitrogen. Nickel-Charged Chelating Sepharose™ Fast Flow was obtained from GE Healthcare. Novapak® Silica Column was purchased from Waters and HPLC hexane was from J.T. Baker-Serviquimia.

ALDH structures and volume measurement of substrate-binding pockets

Human ALDH1A1 (PDB code 4WB9), ALDH1A2 (PDB code 6ALJ) and ALDH1A3 (PDB code 5FHZ; monomer A) crystallographic structures (Table S1) were used for the study of the substrate-binding pocket. The POVME algorithm [20,21] was used to measure the volume of the substrate-binding pockets of the three ALDH1A enzymes. The first step was to define an inclusion region that entirely encompasses all the binding-pocket conformations of the trajectory. Then, a field of equidistant points was generated. After that, points that were near to receptor atoms were removed, leaving those points that were likely to be located within the binding pocket itself. Finally, the patches of points that were not contiguous with the primary binding pocket were removed. The volume of the cofactor-binding pocket was not taken into account in this study, due to the fact that it is a highly conserved region.

Cloning of human ALDH cDNAs

We identified the human genomic clones in GenBank™ (clone ID: 2988388, 4826743 and 6208628 for *ALDH1A1, ALDH1A2* and *ALDH1A3,* respectively). The full-length human *ALDH1A1* and *ALDH1A3* cDNAs were PCR amplified from the pOTB7 vector while human *ALDH1A2* cDNA was amplified from the pBluescriptR vector. ALDH cDNAs were subcloned into pET-30 Xa/LIC using the following primers: 1A1 forward primer (5'-GGTATTGAGGGTCGCATGTCATCCTCAGGCACG-3') and 1A1 reverse primer (5'-AGAGGAGAGTTAGAGCCTTATGAGTTCTTCTGAGAGATTTTCA-3'); 1A2 forward primer (5'-GGTATTGAGGGTCGCATGACTTCCAGCAAGATAGAGATGC-3') and 1A2 reverse primer (5'-AGAGGAGAGTTAGAGCCTTAGGAGTTCTTCTGGGGG-3'); 1A3 forward primer (5'-GGTATTGAGGGTCGCATGGCCACCGCTAACGG-3') and 1A3 reverse primer (5'-AGAGGAGAGTTAGAGCCTCAGGGGTTCTTGTCGCC-3'). Primer sequence extensions required for pET-30 Xa/LIC compatibility are underlined. The PCR steps were 95°C for 2 min, followed by 34 cycles of 95°C for 20 s, 60°C for 10 s, and 70°C for 20 s per kilobase. The resulting construct was transformed into *E. coli* DH5α and cloning was verified by sequencing. Later it was transformed into *E. coli* BL21(DE3)pLys for protein expression.

Construction and cloning of full-length ALDH1A2 cDNA

Four ALDH1A2 isoforms may be generated from alternative splicing of the human *ALDH1A2* gene (isoform 1 corresponds to the full-length enzyme). The difference between isoforms 1 and 2 is the deletion of 38 amino acids, located at residues 288-326 (nucleotides 666-826), in the isoform 2. The commercial source only provided the cDNA coding for the isoform 2 and, for this reason, a method was devised to obtain the full-length cDNA. The complete cDNA encoding isoform 1 of human ALDH1A2

was produced by three independent PCR amplifications. The first PCR used the forward primer referred previously (5'-CGCGGATCCATGACTTCCAGCAAGATAGAGATGC-3') and an internal reverse primer (5'-CCAATGTGAGAAGCTATTGCTGCCCCAGCCGTTGGCCCATATCCTGGCAAAATATTGATC CCGGGAGGAAGCCAGCCTCCTTGATGAGGGCTCC-3') including nucleotides 666-765. The second PCR used an internal forward primer (5'-ATGGGCCAACGGCTGGGGCAGCAATAGCTTCTCACATTGGCATAGACAAGATTGCATTC ACAGGGTCTACTGAGGTTGGAAAGCTTATCCAAGAAGC-3') including nucleotides 726-826 and a reverse primer aforementioned (5'-TGAATTCTTAGGAGTTCTTCTGGGGG-3'). The primer design involved an overlapping region including 39 nucleotides. This fact is essential to ensure the amplification of the ALDH1A2 complete sequence. Finally, a third PCR allowed obtaining the complete sequence from the two fragments amplified previously, by using outer forward and reverse primers and under the same conditions as those for isoform 2 as detailed above. The resulting construct was transformed and verified by sequencing as described above.

Site-directed mutagenesis

Human ALDH1A1 and ALDH1A2 were mutated using the QuickChange Lightning Site-Directed Mutagenesis Kit (Agilent). For this purpose, the following primers were used: For ALDH1A1, ALDH1A1 forward primer (5'-GGTGGAAAACCCTATTCCAATGC-3') and ALDH1A1 reverse primer (5'-GCATTGGAATAGGGTTTTCCACC-3'). For ALDH1A2, ALDH1A2 forward primer (5'-GGATCAATTGTTACGGTGTCGTATCTGCCCAGAGCCCC-3') and ALDH1A2 reverse primer (5'-GGGGCTCTGGGCAGATACGACACCGTAACAATTGATCC-3'). The underlined regions correspond to the mutated nucleotides. The PCR steps were 95°C for 2 min, followed by 18 cycles of 95°C for 20 s, 50°C (ALDH1A1) or 60°C (ALDH1A2) for 10 s, and 68°C for 30s/kb. After that, the PCR products were digested with *Dpn*I to eliminate the parental DNA. The resulting constructs were transformed and verified by sequencing as described above.

Protein expression and purification

Human ALDHs were expressed from the pET-30Xa/LIC constructs, which allowed protein expression with an N-terminal $(His)_6$ tag. For protein expression, transformed *E. coli* BL21(DE3)pLys cells were grown in 25 mL LB medium in the presence of 50 μg/mL kanamycin and 34 μg/mL chloramphenicol at 37°C for 16 h. This culture was used to inoculate 1 L of 2xYT medium in the presence of 50 μg/mL kanamycin and 34 μg/mL chloramphenicol and cells were incubated at 22°C until an $O.D_{595}$ = 0.8 was reached. Then protein expression was induced by the addition of 1 mM IPTG and cells were grown at 22°C for 16 h. To analyze protein expression, 1-mL aliquots were collected before and after induction. After 16-h incubation, cells were collected by centrifugation at 12400 x g (8000 rpm, Beckman JLA-14 rotor) and 4°C for 10 min. The resulting pellet was resuspended in 30 mL of Bind Buffer (20 mM Tris/HCl, 0.5 M NaCl, 5 mM imidazole, pH 8.0) per liter of culture. Cell lysis was performed by freezing (16 h at −20°C) and thawing at room temperature. Then, 1 mg/mL lysozyme, 20 μg/mL DNase, 1% Triton-X100, 1 mM protease inhibitor PMSF and 5 mM DTT were added, followed by two cycles of sonication. The cellular extract was centrifuged at 15000 x g (12500 rpm, Beckman JA-25.5 rotor) and 4°C for 20 min. The soluble fraction was collected and filtered. Protein purification was performed by affinity chromatography on a nickel-charged chelating Sepharose™ Fast Flow 5-mL column (His Trap™ column), which specifically binds the protein due to the His tag, using an ÄKTA™ FPLC purification system. Then, protein was eluted by applying an increasing step-wise concentration (5, 60, 100, and 250 mM) of imidazole in 20 mM Tris/HCl and 0.5 M NaCl, pH 8.0. The enzyme eluted with 250 mM imidazole and imidazole was removed through a PD-10 column (*Gel filtration-desalting column*). Protein including the His tag was stored at −80°C in 20 mM Tris/HCl and 0.5 M NaCl, pH 8.0. To follow the purification process, 1-mL aliquots were collected from each step and analyzed by SDS-PAGE.

Fluorescence assay for the dehydrogenase activity and determination of the kinetic constants

Dehydrogenase activity with non-retinoid substrates was monitored using a fluorimeter (Cary Eclipse Varian) at 25°C. ALDH1A1 and ALDH1A2 were assayed in 50 mM HEPES, 0.5 mM EDTA, pH 8.0, 0.5 mM DTT, while ALDH1A3 assays were performed in 50 mM HEPES, pH 8.0, 30 mM $MgCl_2$, 5 mM DTT. The enzyme concentration in the kinetic assays was in the low nM range. NAD^+ concentration was 500 μM. The enzymatic reaction was initiated by the addition of aldehyde substrate, dissolved in the corresponding reaction buffer. Fluorescence of NADH was followed at 460 nm with excitation at 340 nm and spectral bandwidth of 10 nm. The reaction mixture also contained 5 μM NADH as an internal standard to obtain absolute reaction rates, which were calculated according to the equation: $v = (dF/dt)\cdot(C_{st}/F_{st})$, where C_{st} is the standard NADH concentration, F_{st} the standard fluorescence and dF/dt the slope of the time dependent fluorescence [22]. Assays without enzyme were carried out as controls. The initial velocities were measured in duplicated with at least five different substrate concentrations. The kinetic constants were calculated using the non-linear regression program GraFit 5.0 (Eritacus software), and expressed as the mean ± SEM from three independent determinations.

Effect of magnesium ions

Enzymatic activity was measured fluorimetrically in 50 mM HEPES, 0.5 mM DTT, pH 8.0 (for ALDH1A1 and ALDH1A2) or in 50 mM HEPES, 5 mM DTT, pH 8.0 (for ALDH1A3), in the presence of 500 μM NAD^+. To each reaction mixture, the corresponding amount of $MgCl_2$ was added (0-400 μM final concentration, for ALDH1A1 and ALDH1A2; 0-50 mM, for ALDH1A3). Reaction was initiated with the addition of hexanal as a substrate. Reaction mixtures without enzyme served as controls. The initial rates were measured in duplicate and plotted as the percentage of remaining enzymatic activity.

HPLC-based assay for the dehydrogenase activity with retinoids

To evaluate the retinaldehyde dehydrogenase activity of human ALDHs, an HPLC-based method was used for the first time to determine the *in vitro* ALDH activity, as follows. Stock solutions of retinoid substrates were prepared in ethanol. Working stock solutions of retinoids were prepared by 10-min sonication in presence of equimolar delipidated BSA. The actual concentration of solubilized retinoid was determined based on the corresponding molar absorption coefficient: ε_{370} = 29500 $M^{-1}\cdot cm^{-1}$ for all-*trans*-retinaldehyde and ε_{364} = 26700 $M^{-1}\cdot cm^{-1}$ for 9-*cis*-retinaldehyde. The enzymatic reaction was carried out for 15 min at 37°C in a final volume of 500 μL, using the same buffer as that for the fluorimetric assays (see above) and a saturating concentration of cofactor (500 μM NAD^+). The reaction was stopped by addition of 1 mL cold hexane/dioxane/isopropanol (50:5:1, v/v), and retinoids were extracted by a two-step procedure with the same solvent mixture [23,24]. The aqueous phase was removed and the organic phase was evaporated under a N_2 stream. Retinoids were dissolved in 200 μL of hexane and analyzed by a modification of a published method [25], as follows. Retinoids were separated by HPLC on a NovaPak® Silica 4 μm, 3.9 x 150 mm in hexane/methyl-*tert*-butyl ether (96:4, v/v) mobile phase, at a flow rate of 2 mL/min, using a Waters Alliance 2695 HPLC. Elution was monitored at 370 and 364 nm for all-*trans*- and 9-*cis*-retinaldehyde isomers respectively, and at 350 and 341 nm for the corresponding all-*trans*- and 9-*cis*-RA compounds, with a Waters 2996 photodiode array detector. Quantification of retinoids was performed by interpolation of HPLC peak areas into a calibration curve of known retinoid concentrations. Kinetic constants were calculated as indicated above.

Spectrophotometric assay for esterase activity

Enzymes were assayed for esterase activity in a Cary 400 Bio (Varian) spectrophotometer. Stock solutions of 10 mM *p*-nitrophenyl acetate were prepared with ethanol as a solvent to minimize spontaneous hydrolysis, always ensuring that the final ethanol concentration in the assay did not exceed 1% (v/v) [26]. Reactions were performed at 25°C in 50 mM HEPES, pH 7.5, 0.5 mM DTT, 500 μM NAD^+. The reactions were initiated by the addition of *p*-nitrophenyl acetate to the enzyme solutions in 1-mL cuvettes. The formation of *p*-nitrophenol was followed at 400 nm (ε_{400} = 12800 $M^{-1}\cdot cm^{-1}$) [27] to monitor esterase activity, which was corrected for the non-enzymatic hydrolysis of *p*-nitrophenyl acetate. Kinetic constants were calculated as indicated above.

Results

Comparison of the substrate-binding pocket of ALDH1A enzymes

The recent availability of the crystal structures of the holo-ALDH1A enzymes (Table S1) has allowed us to perform for the first time a comparative topological analysis of the substrate-binding pockets of the three human ALDH1A holoenzymes. The three structures are well superimposed (pairwise RMSD values ranged from 0.429 to 0.659 Å) but they are not fully equivalent, as they correspond to different complexes with oxidized/reduced cofactor and with or without inhibitor or product molecule: ALDH1A1 with NADH, ALDH1A2 with NAD^+ and WIN18,446 inhibitor, and ALDH1A3 with NAD^+ and RA. The active-site nucleophile Cys residue and the cofactor molecule could be well superimposed in the three structures. The contracted or open conformations of NAD^+ in the ALDH1A2 and ALDH1A3 ternary complexes were similar to that of NADH in the ALDH1A1 binary complex [8–10]. The previously deposited ALDH1A2-NAD^+ structure (PDB code 4X2Q) was incomplete as it included an unstructured segment between Asn475 and Met495 (corresponding to a loop region, Gly458-Leu478 in ALDH1A1 and Asn469-Leu489 in ALDH1A3), located between the β18 sheet and α13 helix, and it contained only a fragment of NAD^+ presumably due to incomplete ligand occupancy. Overall, ALDH1A1, ALDH1A2 and ALDH1A3 exhibited similar topologies and decreasing atomic volumes in their substrate-binding pockets, i.e. 534, 387 and 357 $Å^3$, respectively (Figure 1). In the three structures, invariant Trp178 (ALDH1A1 numbering) and residue 460 (Val/Leu) define a bottle neck in the middle part of the pocket. In ALDH1A2, a new subpocket, lined by loop residues Leu477, Asn478 and Ala479, is observed near the solvent exposed area, which could be generated by inhibitor binding [10]. Table 1 lists the residues lining the substrate-binding pocket in the ALDH1A enzymes. In an entire 20-residue ALDH1A2 segment (residues 475-495) it is possible to identify six residue changes in comparison with ALDH1A1: Asn475, Ala476, Leu477, Asn478, Ser481 and Met495, located in ALDH1A2, are replaced by Gly458, Val459, Val460, Ser461, Cys464 and Leu478, respectively, in the ALDH1A1 isozyme. However, only three residues, Asn478, Ser481 and Met495, are replaced by Tyr472, Ala475 and Leu489, respectively, in ALDH1A3. The first four contiguous residues, 475 to 478, were chosen to construct an ALDH1A2 mutant in order to mimic ALDH1A1 architecture. Residue 114, located in the NAD^+-binding domain just before helix α3 (residues 115-136), which defines one side of the pocket, is Leu in ALDH1A1 but Pro in ALDH1A2 and ALDH1A3. Leu is typically a residue with a high α-helix-forming propensity while Pro could be helix breaker depending on its environment. Thus, the change L114P was selected for a mutant ALDH1A1 to make this part of the substrate-binding pocket more similar to that of ALDH1A2.

Effect of Mg^{2+} ions on the dehydrogenase activity of human ALDH1A enzymes

The inhibitory effect of Mg^{2+} ions and other divalent cations on ALDH1A1 (formerly referred as cytosolic ALDH or ALDH1) from human and other species had been reported, using acetaldehyde or propanal as substrates [28]. Here we tested the Mg^{2+} effect by using hexanal, since it is the substrate with the highest catalytic efficiency (Table 2). When examining the effect of Mg^{2+} ions on the dehydrogenase activity of ALDH1A enzymes, a quite distinct pattern was observed for each of the three enzymes (Figure 2). ALDH1A1 and ALDH1A2 activity decreased with increasing concentrations of Mg^{2+}, while ALDH1A3 activity increased. ALDH1A1 was extremely sensitive to divalent cations and thus it lost more than 80% of its activity in the presence of 50 μM $MgCl_2$. ALDH1A2 activity exhibited a less pronounced inhibitory pattern, losing only up to 40% of its activity when the Mg^{2+} concentration reached 300 μM. The strongest inhibition of the ALDH1A2 activity by magnesium was observed at 250-400 μM Mg^{2+}, which is close to the cofactor concentration used in the assay. This effect could have been related to the formation of an NAD^+-Mg^{2+} complex that might have resulted in a slower NAD^+-enzyme binding. If that had been the case, the decreased ALDH1A2 activity would have been caused not just by the magnesium effect on the enzyme but rather by a decreased NAD^+ availability. To discard this possibility, a magnesium titration experiment for ALDH1A2 was performed at different concentrations of NAD^+

(250, 500 and 1000 μM). The results (not shown) indicate that the effect of magnesium ions on ALDH1A2 is not dependent on the cofactor concentration used.

In contrast, Mg^{2+} caused an opposite effect on ALDH1A3, since the enzyme activity increased as much as three-fold at 5 mM $MgCl_2$ and four-fold when the $MgCl_2$ concentration was raised up to 30 mM. Thus, thereafter, ALDH1A1 and ALDH1A2 were routinely assayed in the absence of Mg^{2+} and with 0.5 mM EDTA, while ALDH1A3 assays were performed in 30 mM $MgCl_2$ without any chelating agent.

Kinetic characterization of human ALDH1A enzymes with non-retinoid aldehydes

A wide survey of the substrate specificity of the ALDH1A forms was performed by determining their kinetic parameters with cofactor NAD^+ and a variety of aldehyde compounds, namely alkanals and alkenals (Figure S1), including lipid peroxidation products (Table 2). In terms of k_{cat}/K_m, ALDH1A3 exhibits the lowest values for all substrates, suggesting a moderate role in the physiological oxidation of these aldehydes. The k_{cat}/K_m values of ALDH1A1 and ALDH1A2 indicate a potentially major role in the transformation of these substrates with slightly different substrate specificity. Among the three enzymes, ALDH1A1 showed low K_m and k_{cat} values for most of the substrates, while ALDH1A2 exhibited the highest k_{cat} value. With citral (3,7-dimethyl-2,6-octadienal), which shares its isoprenoid nature with retinoids, ALDH1A1 showed the lowest K_m (0.085 μM) and k_{cat} (1 min^{-1}) values. In fact, citral is a slow substrate but with high affinity and has often been used as a competitive inhibitor of ALDH1A1 [29]. The enzyme concentration used in the kinetic assays was strictly monitored and kept within the low nM range. Specifically, the final enzyme concentration with low-K_m substrates, such as citral, was 1.1 nM. In those cases, the substrate concentration was at least 80-fold higher than the enzyme concentration and the kinetics could be fitted using the standard Michaelis-Menten equation. Hexanal was the best substrate for the three enzymes. Saturated aldehydes (hexanal and octanal), exhibited higher catalytic efficiency than their unsaturated counterparts (*trans*-2-hexenal and *trans*-2-octenal). Regarding the aldehydes derived from lipid peroxidation, 4-hydroxy-2-hexenal and 4-hydroxy-2-nonenal, the activity of the ALDH1A forms is comparable to that of other ALDHs [30], suggesting a shared role in the oxidation of these physiological substrates. ALDH1A2 is the best enzyme for 4-hydroxy-2-nonenal in terms of k_{cat}/K_m. Finally, the three enzymes showed esterase activity in the presence of NAD^+, being ALDH1A2 the form with the highest catalytic efficiency.

Kinetic characterization of human ALDH1A enzymes with retinaldehyde

The activity with retinoid molecules was determined by using an HPLC-based method to analyze the substrates and reaction products (Figure 3). Different solvent mixtures were tested to optimize retinoid recovery after the enzymatic reaction: hexane/dioxane/isopropanol (50:5:1) [23], hexane/ethyl acetate (9:1) [31], hexane [32], and a two step acid-base extraction [25]. From the evaluated methods, the extraction with hexane/dioxane/isopropanol (50:5:1) was chosen since it was the most efficient in recovering retinaldehyde and RA from the activity buffer, with a yield near 100%. To our knowledge, this is the first time that this extraction method and the use of BSA in the activity assay have been applied to determine the enzymatic activity of ALDH1A enzymes with retinoids. In all cases, concomitant retinaldehyde consumption and RA production, under steady state conditions, could be measured. The three ALDH1A enzymes followed Michaelis-Menten kinetics with both retinaldehyde isomers (exemplified in Figure 4 for all-*trans*-retinaldehyde), with K_m values in the micromolar range. Table 3 shows the kinetic constants of the ALDH1A enzymes with all-*trans* and 9-*cis* retinaldehyde. ALDH1A enzymes were fairly active with the two isomers, with k_{cat}/K_m values within the upper range of most aldehyde substrates (Table 2). The three forms exhibited lower K_m values for 9-*cis*-retinaldehyde than for all-*trans*-retinaldehyde. Regarding the k_{cat} values, they were higher for the all-*trans* isomer (ALDH1A2 and ALDH1A3) or similar for the two isomers (ALDH1A1). This results in that for the most physiological isomer, all-*trans*-retinaldehyde, ALDH1A3 is the best enzyme in terms of k_{cat}/K_m, followed by ALDH1A2. Conversely, the k_{cat} value and the catalytic efficiency of ALDH1A3 is lowest for the 9-*cis* isomer. The all-*trans*-retinaldehyde specificity of ALDH1A3 within the ALDH1A subfamily is a

remarkable feature given the fact that this enzyme exhibited the lowest k_{cat}/K_m values with all examined non-retinoid substrates (Table 2).

Kinetic characterization of mutant ALDH1A enzymes

In order to test the role of specific residues in determining the kinetic properties of ALDH1A enzymes, we performed site-directed mutagenesis of selected residues based on structural differences. The substitution L114P, defining one site of the substrate-binding pocket, was selected for a mutant ALDH1A1 to make this part of the structure more similar to that of ALDH1A2. Concurrently, in ALDH1A2, four contiguous residue changes, N475G, A476V, L477V and N478S (corresponding to the region with the loop) were made to mimic the structure of ALDH1A1. The cofactor NAD^+, three typical aldehyde substrates and both retinaldehyde isomers were tested. Table 4 shows the comparison of the kinetic constants of the wild-type enzymes and their corresponding mutants. The cofactor kinetic constants for the mutants showed only moderate variations as compared with those of the wild-type enzymes, suggesting that NAD^+ binding was not significantly affected by the mutations. For the ALDH1A1 L114P mutant, there was a significant increase (50-100 fold) in the K_m values for hexanal and citral. This caused the catalytic efficiency for these substrates, especially citral, to be more similar to that of wild-type ALDH1A2. Conversely, the major effect observed with the mutant ALDH1A2 was a 50-fold decrease in the K_m value for citral. This decline, together with the 5-fold decrease in the k_{cat} value, implies that the catalytic efficiency of mutant ALDH1A2 for citral was closer to that observed in wild-type ALDH1A1. Regarding kinetics with retinaldehyde isomers, the mutants did not show significant differences with the respective wild-type forms, suggesting that the mutated residues are not critical for retinaldehyde specificity.

Discussion

The current comparative study on the structure and kinetic properties of ALDH1A enzymes has a genuine interest stemming out from the fact that they have a proven physiological role in RA synthesis, both in adult organisms as well as in embryos. The enzymes from the ALDH1A subfamily are NAD^+-dependent cytosolic retinaldehyde dehydrogenases involved in the last metabolic step of RA biosynthesis. Their essential role in RA homeostasis and signaling has been widely shown by knockout mice studies [33] and, recently, by their involvement in genetic diseases [34,35].

For more than a decade, the only known ALDH1A structures were those of the sheep ALDH1 (ALDH1A1) [36] and rat RALDH2 (ALDH1A2) [37]. The recent reports on the X-ray coordinates of the human ALDH1A enzymes (ALDH1A1, ALDH1A2 and ALDH1A3) have enabled for the first time the close comparison of their three-dimensional structures. As expected from their overall 70% of amino acid sequence identity, the three enzymes exhibit a high degree of structural similarity, but demonstrate significant differences within their substrate-binding pockets. Although ALDH1A1, ALDH1A2 and ALDH1A3 show similar topologies, they have decreasing enclosed volumes in their substrate-binding pockets, with ALDH1A2 showing a distinct subpocket associated to its 475-495 loop. Most residues in the active-site channel are identical, 9 out of 12 residues, between ALDH1A2 and ALDH1A3, while only 5 residues are shared with ALDH1A1 (Table 1). Most residue differences in ALDH1A1 are concentrated in the loop region. Ile132, Gly136, Arg139, Thr140, Trp189, Leu471 and Ala473 establish van der Waals interactions with RA in ALDH1A3 [9]. Three residues, 125, 460 and 304 (ALDH1A1 numbering), located at the channel mouth, bottom and neck, respectively, were previously pointed as important for the substrate-binding pocket to determine the ability to harbor either small or large substrates in ALDH1 *vs.* ALDH2 enzymes [38]. Importantly, residue 304 interacts with RA in ALDH1A3 [9]. Residue 125 remains an invariable Gly in the three ALDH1A enzymes and residue 460 shows a conservative substitution (a hydrophobic Val or Leu), but residue 304 involves a significant exchange (Ile/Thr) (Table 1). Residues 458 (Gly in ALDH1A1, and Asn in ALDH1A2 and ALDH1A3) and 461 (Ser in ALDH1A1, Asn in ALDH1A2, and Tyr472 in ALDH1A3) display even more remarkable substitutions. Such variability may be important in shaping the enclosed volume and topology of the

substrate-binding site, and in modulating substrate specificity and inhibitor selectivity, as previously noticed for residue 458 [39] .

The study of the substrate-binding pocket aims to tackle the question whether specific residue differences could be affecting enzymatic activity. Previously reported kinetic parameters were obtained using a plethora of experimental conditions, including buffer type, pH, use of Mg^{2+} ions, chelating and reducing agents, and solubilizing agents [9,18,19,40–42]. Here, for the first time, the activity of the three human enzymes was characterized side-by-side under similar experimental conditions, making possible the most thorough comparative study reported so far. In addition, the data obtained in this work could be checked against with those of previous reports. For human ALDH1A1, our results with typical aldehydes (*trans*-2-hexenal and citral) are similar to the few data available [22]. For human ALDH1A3, the obtained results are comparable with those of rat ALDH1A3 [43].

Given the relevance of ALDH1A enzymes to RA production, we also investigated the contribution of each enzyme to retinaldehyde oxidation. In order to determine retinaldehyde activity, we followed, for the first time in ALDH studies, a well-proven strategy for retinoid solubilization (with fatty acid-free bovine serum albumin) [44] and an improved HPLC-based method to unequivocally monitor substrate disappearance and product formation using an end-point assay. This combined approach allowed us to work with a very low concentration of substrate and in the absence of any added detergent or organic solvent. This is not feasible with other methodologies, such as UV-vis spectrophotometry, which have been classically used in previous studies [9,19,40,42] The three ALDH1A enzymes are active with the two physiological retinaldehyde isomers, meaning that their distinct active-sites allow binding of retinoid molecules, as directly observed by RA binding in the ALDH1A3 crystallographic structure [9]. Thus, residues Asn469, Leu471 and Ala473 have been pointed out as making hydrophobic contacts with the retinoid molecule in the ALDH1A3 structure. Only Ala473 is strictly conserved in the three ALDH1A enzymes, while Asn469 and Leu471 are conserved in ALDH1A2 (Table 4). Despite of exchanging some of these residues important for retinoid binding, between ALDH1A and ALDH1A2, no major effect was observed in the kinetic constants for all-*trans*-retinaldehyde, supporting the idea that some variability is allowed within the loop residues without compromising retinoid binding. This variability may be linked with the retinoid isomer specificity. In this regard, ALDH1A3 is truly specific for all-*trans*-retinaldehyde, as its catalytic efficiency with this substrate is 80-fold higher than with 9-*cis*-retinaldehyde. The unfavorable kinetics of ALDH1A3 with the 9-*cis* isomer may be explained by steric hindrance due to the smaller volume of the substrate-binding pocket, compared to those of ALDH1A1 and ALDH1A2. ALDH1A3 appears to be the most conspicuous all-*trans* retinaldehyde dehydrogenase since it exhibits the highest k_{cat}/K_m values among the ALDH1A forms, and this substrate is the most efficiently catalyzed by ALDH3 other than hexanal. Citral specificity for ALDH1A1 could linked in part to residue 114 (ALDH1A1 L114P mutant) and to residues 458 to 461 (ALDH1A2 mutant).

Previously, the enzymatic activity of these proteins with retinoids had been explored. Thus, the kinetic constants with retinaldehyde have been reported for ALDH1A1 [18,40], ALDH1A2 [19] and ALDH1A3 [35]. However, a single study with a common methodology and under the same experimental conditions was lacking. As for the orthologous murine enzymes, the retinoid activity of rat ALDH1A1 was characterized [45] and the kinetic constants of retinoids were determined for rat [40,46] and mouse [47] ALDH1A2. Similarly to the human orthologous proteins, mouse and rat ALDH1A2 also catalyze the oxidation of both all-*trans*- and 9-*cis*-retinaldehyde. The enzymatic activity of mouse ALDH1A3 was characterized with several aldehydes [43] and, like the human enzyme, ALDH1A3 is highly specific for all-*trans* retinaldehyde [48].

ALDH1A1 is highly expressed in the dorsal retina of mouse embryo [49] in epithelial tissues of adult mice and Xenopus [50], and in the stomach and small intestine of adult rats [51]. The knockout of ALDH1A1 did not severely affect the morphology of the retina [52], indicating that other ALDH1A enzymes might redundantly share the function of ALDH1A1. In addition, ALDH1A1 is able to induce the differentiation of hematopoietic and neural stem cells and has been widely used as a marker of cancer stem cells. ALDH1A1 is also involved in chemoresistance, specifically against cyclophosphamide treatment [53]. ALDH1A2 is the major form involved in early embryonic and in cardiac development

[54]. Non-conservative ALDH1A2 mutations are associated with rare cases of human congenital heart disease [55]. Interestingly, the most severe form of RA signaling deprivation is the genetic deficiency of ALDH1A2, which is lethal at 8.5 days post-fertilization in mice [56]. The phenotypes of ALDH1A2 knock-out mice include severely impaired segmentation of rhombomeres, altered homeobox gene expression pattern, and defective neural crest cell migration [57]. In adults, ALDH1A2 is involved in spermatogenesis, which is the basis of inhibitor development intended as anticonceptive drugs [10,58]. ALDH1A3 is expressed in the ventral retina across various species [59,60] and is essential for the development of the central nervous system and the morphogenesis of anterior head structures. Mice lacking ALDH1A3 have severe defects in nasal and ocular development and are neonatal lethal, due to respiratory tract obstruction in the nasal region, but they can be rescued by RA administration during pregnancy [61]. Recently, several groups have reported that ALDH1A3 is associated with anophthalmia/microphthalmia in humans [35].

The three ALDH1A enzymes are therefore essential for embryonic and adult life because of their role in RA synthesis. The low micromolar range K_m values with all-*trans*-retinal, the most physiologically relevant isomer, and high catalytic efficiency, common for human and murine enzymes, are suitable for this role. The higher efficiency of ALDH1A2 and ALADH1A3 correlates with the stronger effects of the respective knock-outs.

Interestingly, the effect of Mg^{2+} ions, using hexanal as a substrate, is quite distinct for each ALDH1A enzyme. While ALDH1A1 activity is severely inhibited, ALDH1A2 activity is moderately decreased and, in contrast, ALDH1A3 is significantly increased. The relevant Mg^{2+} concentrations are within the range of the physiological intracellular concentrations of free Mg^{2+} (0.5-1 mM) [62]. The inhibitory effect of Mg^{2+} ions has been described previously for ALDH1A1 and seems to be related to coenzyme dissociation [27]. Magnesium stabilizes cofactor by interacting with the pyrophosphate group, tightening NADH binding and making cofactor dissociation slower. The decrease in activity of ALDH1A1 with Mg^{2+} is therefore evidence that coenzyme dissociation is the rate-limiting step, as previously proposed [27]. The rate-limiting steps for ALDH1A2 and ALDH1A3 are not known at this time. The moderate inhibitory effect of Mg^{2+} ions on ALDH1A2 may support a limiting step similar to that of ALDH1A1, the coenzyme dissociation. The activating effect on ALDH1A3 is consistent with what was previously found for mouse ALDH1A3 [43]. This activation is reminiscent of the reported effect on the mitochondrial ALDH2 form, where Mg^{2+} ions enhanced deacylation, which is the rate-limiting step for this enzyme [28]. The similar activation by Mg^{2+} ions observed here for ALDH1A3 gives support for deacylation being the rate-limiting step for ALDH1A3. Thus, our results are consistent with cofactor dissociation being the rate-limiting step for ALDH1A1 and ALDH1A2, and deacylation for ALDH1A3, when hexanal is used as a substrate.

In summary, it can be concluded that the three enzymes, ALDH1A1, ALDH1A2 and ALDH1A3, share a similar substrate-binding pocket topology (with some scattered point residue differences and decreasing enclosed volumes) and common substrate specificity (with some notable singularities). Some kinetic features, such as citral specificity for ALDH1A1, have been linked in part to residue 114 and to residues 458 to 461 of the loop located between β18 sheet and α13 helix. Specificity of ALDH1A3 for all-*trans*- *versus* the 9-*cis*- isomer of retinaldehyde might be related to the smaller volume of its substrate-binding pocket. In addition to the known role of ALDH1A enzymes in the conversion of retinaldehyde to RA, they all may participate in aldehyde detoxification and, specifically, in the elimination of lipid peroxidation products (*trans*-2-hexenal, 4-hydroxy-2-hexenal and 4-hydroxy-2-nonenal). Both types of activity are relevant to the physiological roles of ALDH1A enzymes, such embryogenesis or stemness maintenance in cancer stem cells.

Acknowledgements: R.P. was a recipient of a PIF predoctoral fellowship from Universitat Autònoma de Barcelona.

Funding Information: This work has been funded by the Spanish Ministerio de Economía y Competitividad (BIO2016-78057).

References

[1] S.A. Marchitti, C. Brocker, D. Stagos, V. Vasiliou, Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily, Expert Opin Drug Metab Toxicol. 4 (2008) 697–720.

[2] V. Vasiliou, A. Pappa, T. Estey, Role of human aldehyde dehydrogenases in endobiotic and xenobiotic metabolism, Drug Metab Rev. 36 (2004) 279–299.

[3] A. Pappa, C. Chen, Y. Koutalos, A.J. Townsend, V. Vasiliou, Aldh3a1 protects human corneal epithelial cells from ultraviolet- and 4-hydroxy-2-nonenal-induced oxidative damage, Free Radic Biol Med. 34 (2003) 1178–1189.

[4] N. Lassen, A. Pappa, W.J. Black, J. V Jester, B.J. Day, E. Min, et al., Antioxidant function of corneal ALDH3A1 in cultured stromal fibroblasts, Free Radic Biol Med. 41 (2006) 1459–1469.

[5] T. Estey, M. Cantore, P.A. Weston, J.F. Carpenter, J.M. Petrash, V. Vasiliou, Mechanisms involved in the protection of UV-induced protein inactivation by the corneal crystallin ALDH3A1, J Biol Chem. 282 (2007) 4382–4392.

[6] N. Lassen, J.B. Bateman, T. Estey, J.R. Kuszak, D.W. Nees, J. Piatigorsky, et al., Multiple and additive functions of ALDH3A1 and ALDH1A1: cataract phenotype and ocular oxidative damage in Aldh3a1(-/-)/Aldh1a1(-/-) knock-out mice, J Biol Chem. 282 (2007) 25668–25676.

[7] L. Uma, J. Hariharan, Y. Sharma, D. Balasubramanian, Corneal aldehyde dehydrogenase displays antioxidant properties, Exp Eye Res. 63 (1996) 117–120.

[8] C.A. Morgan, T.D. Hurley, Development of a high-throughput in vitro assay to identify selective inhibitors for human ALDH1A1, Chem Biol Interact. 234 (2015) 29–37.

[9] A. Moretti, J. Li, S. Donini, R.W. Sobol, M. Rizzi, S. Garavaglia, Crystal structure of human aldehyde dehydrogenase 1A3 complexed with NAD+ and retinoic acid, Sci. Rep. 6 (2016) 35710.

[10] Y. Chen, J.-Y. Zhu, K.H. Hong, D.C. Mikles, G.I. Georg, A.S. Goldstein, et al., Structural basis of ALDH1A2 inhibition by irreversible and reversible small molecule inhibitors, ACS Chem. Biol. 13 (2018) 582–590.

[11] J. Farrés, T.T. Wang, S.J. Cunningham, H. Weiner, Investigation of the active site cysteine residue of rat liver mitochondrial aldehyde dehydrogenase by site-directed mutagenesis., Biochemistry. 34 (1995) 2592–2598.

[12] C. Canestro, J.M. Catchen, A. Rodriguez-Mari, H. Yokoi, J.H. Postlethwait, Consequences of lineage-specific gene loss on functional evolution of surviving paralogs: ALDH1A and retinoic acid signaling in vertebrate genomes, PLoS Genet. 5 (2009) e1000496.

[13] F.X. Ruiz, S. Porte, X. Pares, J. Farres, Biological role of aldo-keto reductases in retinoic Acid biosynthesis and signaling, Front Pharmacol. 3 (2012) 58.

[14] R.M. Connolly, N.K. Nguyen, S. Sukumar, Molecular pathways: current role and future directions of the retinoic acid pathway in cancer prevention and treatment, Clin Cancer Res. 19 (2013) 1651–1659.

[15] G. Smith, C.R. Wolf, Y.Y. Deeni, R.S. Dawe, A.T. Evans, M.M. Comrie, et al., Cutaneous expression of cytochrome P450 CYP2S1: Individuality in regulation by therapeutic agents for psoriasis and other skin diseases, Lancet. 361 (2003) 1336–1343.

[16] G. Duester, F.A. Mic, A. Molotkov, Cytosolic retinoid dehydrogenases govern ubiquitous metabolism of retinol to retinaldehyde followed by tissue-specific metabolism to retinoic acid, Chem Biol Interact. 143-144 (2003) 201–210.

[17] P. Marcato, C. a. Dean, C. a. Giacomantonio, P.W.K. Lee, Aldehyde dehydrogenase its role as a cancer stem cell marker comes down to the specific isoform, Cell Cycle. 10 (2011) 1378–1384.

[18] P. V Bhat, H. Samaha, Kinetic properties of the human liver cytosolic aldehyde dehydrogenase for retinal isomers, Biochem Pharmacol. 57 (1999) 195–197.

[19] Y. Shabtai, H. Jubran, T. Nassar, J. Hirschberg, A. Fainsod, Kinetic characterization and regulation of the human retinaldehyde dehydrogenase 2 enzyme during production of retinoic acid, Biochem J. 473 (2016) 1423–1431.

[20] J.D. Durrant, C.A. de Oliveira, J.A. McCammon, POVME: an algorithm for measuring binding-pocket volumes, J Mol Graph Model. 29 (2011) 773–776.

[21] J.D. Durrant, L. Votapka, J. Sorensen, R.E. Amaro, POVME 2.0: An Enhanced Tool for Determining Pocket Shape and Volume Characteristics, J Chem Theory Comput. 10 (2014) 5047–5056.

[22] S. Solobodowska, J. Giebultowicz, R. Wolinowska, P. Wroczynski, Contribution of ALDH1A1 isozyme to detoxification of aldehydes present in food products, Acta Pol Pharm. 69 (2012) 1380–1383.

[23] P. McCaffery, J. Evans, O. Koul, A. Volpert, K. Reid, M.D. Ullman, Retinoid quantification by HPLC/MS ^{n}, J. Lipid Res. 43 (2002) 1143–1149.

[24] J.E. Evans, P. McCaffery, HPLC/MS Analysis of Retinoids, in: H. Sun, G.H. Travis (Eds.), Retin. Methods Protoc. Chapter 8, Humana Press, Totowa, NJ, 2010: pp. 149–162.

[25] M.A. Kane, N. Chen, S. Sparks, J.L. Napoli, Quantification of endogenous retinoic acid in limited biological samples by LC/MS/MS, Biochem J. 388 (2005) 363–369.

[26] R.I. Feldman, H. Weiner, Horse liver aldehyde dehydrogenase. II. Kinetics and mechanistic implications of the dehydrogenase and esterase activity, J Biol Chem. 247 (1972) 267–272.

[27] H. Weiner, J.H. Hu, C.G. Sanny, Rate-limiting steps for the esterase and dehydrogenase reaction catalyzed by horse liver aldehyde dehydrogenase, J Biol Chem. 251 (1976) 3853–3855.

[28] K.K. Ho, A. Allali-Hassani, T.D. Hurley, H. Weiner, Differential effects of Mg2+ ions on the individual kinetic steps of human cytosolic and mitochondrial aldehyde dehydrogenases, Biochemistry. 44 (2005) 8022–8029.

[29] V. Koppaka, D.C. Thompson, Y. Chen, M. Ellermann, K.C. Nicolaou, R.O. Juvonen, et al., Aldehyde dehydrogenase inhibitors: a comprehensive review of the pharmacology, mechanism of action, substrate specificity, and clinical application, Pharmacol Rev. 64 (2012) 520–539.

[30] J. Brichac, K.H. Kwok, A. Honzatko, R. Wang, X. Lu, H. Weiner, et al., Enantioselective oxidation of trans-4-hydroxy-2-nonenal is aldehyde dehydrogenase isozyme and Mg2+dependent, Chem. Res. Toxicol. 20 (2007) 887–895.

[31] Y. Yamakoshi, H. Fukasawa, T. Yamauchi, H. Waki, T. Kadowaki, K. Shudo, et al., Determination of Endogenous Levels of Retinoic Acid Isomers in Type II Diabetes Mellitus Patients. Possible Correlation with HbA1c Values, Biol. Pharm. Bull. 25 (2002) 1268–1271.

[32] Y.K. Kim, L. Quadro, HPLC/MSN Analysis of Retinoids, in: H. Sun, G.H. Travis (Eds.), Retin. Methods Protoc. Chapter 15, Humana Press, Totowa, NJ, 2010: pp. 149–162.

[33] S. Kumar, L.L. Sandell, P.A. Trainor, F. Koentgen, G. Duester, Alcohol and aldehyde dehydrogenases: retinoid metabolic effects in mouse knockout models, Biochim Biophys Acta. 1821 (2012) 198–205.

[34] L. Fares-Taie, S. Gerber, N. Chassaing, J. Clayton-Smith, S. Hanein, E. Silva, et al., ALDH1A3 Mutations Cause Recessive Anophthalmia and Microphthalmia, Am J Hum Genet. (2012) 265-270.

[35] A. Mory, F.X. Ruiz, E. Dagan, E.A. Yakovtseva, A. Kurolap, X. Pares, et al., A missense mutation in ALDH1A3 causes isolated microphthalmia/anophthalmia in nine individuals from an inbred Muslim kindred, Eur J Hum Genet. (2014) 419-422.

[36] S.A. Moore, H.M. Baker, T.J. Blythe, K.E. Kitson, T.M. Kitson, E.N. Baker, Sheep liver cytosolic aldehyde dehydrogenase: the structure reveals the basis for the retinal specificity of class 1 aldehyde dehydrogenases., Structure. 6 (1998) 1541–1551.

[37] A.L. Lamb, M.E. Newcomer, The structure of retinal dehydrogenase type II at 2.7 A resolution: implications for retinal specificity., Biochemistry. 38 (1999) 6003–6011.

[38] T.J.P. Sobreira, F. Marlétaz, M. Simões-Costa, D. Schechtman, A.C. Pereira, F. Brunet, et al., Structural shifts of aldehyde dehydrogenase enzymes were instrumental for the early evolution of retinoid-dependent axial patterning in metazoans., Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 226–231.

[39] C.A. Morgan, T.D. Hurley, Characterization of two distinct structural classes of selective aldehyde dehydrogenase 1A1 inhibitors, J Med Chem. 58 (2015) 1964–1975.

[40] R. Bchini, V. Vasiliou, G. Branlant, F. Talfournier, S. Rahuel-Clermont, Retinoic acid biosynthesis catalyzed by retinal dehydrogenases relies on a rate-limiting conformational transition associated with substrate recognition, Chem Biol Interact. 202 (2013) 78–84.

[41] J. Paik, M. Haenisch, C.H. Muller, A.S. Goldstein, S. Arnold, N. Isoherranen, et al., Inhibition of retinoic acid biosynthesis by the bisdichloroacetyldiamine WIN 18,446 markedly suppresses spermatogenesis and alters retinoid metabolism in mice, J. Biol. Chem. 289 (2014) 15104–15117.

[42] A.A. Klyosov, Kinetics and specificity of human liver aldehyde dehydrogenases toward aliphatic, aromatic, and fused polycyclic aldehydes, Biochemistry. 35 (1996) 4457–4467.

[43] C.E. Graham, K. Brocklehurst, R.W. Pickersgill, M.J. Warren, Characterization of retinaldehyde dehydrogenase 3, Biochem J. 394 (2006) 67–75.

[44] N.Y. Kedishvili, W.H. Gough, W.I. Davis, S. Parsons, T.K. Li, W.F. Bosron, Effect of cellular retinol-binding protein on retinol oxidation by human class IV retinol/alcohol dehydrogenase and inhibition by ethanol, Biochem. Biophys. Res. Commun. 249 (1998) 191–196.

[45] J. Labrecque, F. Dumas, A. Lacroix, P. V Bhat, A novel isoenzyme of aldehyde dehydrogenase specifically involved in the biosynthesis of 9-cis and all-trans retinoic acid, Biochem J. 305 (1995) 681–684.

[46] X. Wang, P. Penzes, J.L. Napoli, Cloning of a cDNA encoding an aldehyde dehydrogenase and its expression in Escherichia coli. Recognition of retinal as substrate, J Biol Chem. 271 (1996) 16288–16293.

[47] I. Gagnon, G. Duester, P. V Bhat, Kinetic analysis of mouse retinal dehydrogenase type-2 (RALDH2) for retinal substrates, Biochim Biophys Acta. 1596 (2002) 156–162.

[48] A. Sima, M. Parisotto, S. Mader, P. V Bhat, Kinetic characterization of recombinant mouse retinal dehydrogenase types 3 and 4 for retinal substrates, Biochim Biophys Acta. 1790 (2009) 1660–1664.

[49] R.J. Haselbeck, I. Hoffmann, G. Duester, Distinct functions for Aldh1 and Raldh2 in the control of ligand production for embryonic retinoid signaling pathways, Dev Genet. 25 (1999) 353–364.

[50] H.L. Ang, G. Duester, Retinoic acid biosynthetic enzyme ALDH1 localizes in a subset of retinoid-dependent tissues during xenopus development, Dev Dyn. 215 (1999) 264–272.

[51] A. Frota-Ruchon, M. Marcinkiewicz, P. V Bhat, Localization of retinal dehydrogenase type 1 in the stomach and intestine, Cell Tissue Res. 302 (2000) 397–400.

[52] X. Fan, A. Molotkov, S. Manabe, C.M. Donmoyer, L. Deltour, M.H. Foglio, et al., Targeted disruption of Aldh1a1 (Raldh1) provides evidence for a complex mechanism of retinoic acid synthesis in the developing retina, Mol Cell Biol. 23 (2003) 4637–4648.

[53] N.E. Sladek, Aldehyde dehydrogenase-mediated cellular relative insensitivity to the oxazaphosphorines, Curr Pharm Des. 5 (1999) 607–625.

[54] K. Niederreither, P. McCaffery, U.C. Drager, P. Chambon, P. Dolle, Restricted expression and retinoic acid-induced downregulation of the retinaldehyde dehydrogenase type 2 (RALDH-2) gene during mouse development, Mech Dev. 62 (1997) 67–78.

[55] M. Pavan, V.F. Ruiz, F.A. Silva, T.J. Sobreira, R.M. Cravo, M. Vasconcelos, et al., ALDH1A2 (RALDH2) genetic variation in human congenital heart disease, BMC Med Genet. 10 (2009) 113.

[56] K. Niederreither, V. Subbarayan, P. Dolle, P. Chambon, Embryonic retinoic acid synthesis is essential for early mouse post-implantation development, Nat Genet. 21 (1999) 444–448.

[57] K. Niederreither, J. Vermot, B. Schuhbaur, P. Chambon, P. Dolle, Retinoic acid synthesis and hindbrain patterning in the mouse embryo, Development. 127 (2000) 75–85.

[58] J.K. Amory, C.H. Muller, J.A. Shimshoni, N. Isoherranen, J. Paik, J.S. Moreb, et al., Suppression of spermatogenesis by bisdichloroacetyldiamines is mediated by inhibition of testicular retinoic acid biosynthesis, J. Androl. 32 (2011) 111–119.

[59] R. Suzuki, T. Shintani, H. Sakuta, A. Kato, T. Ohkawara, N. Osumi, et al., Identification of RALDH-3, a novel retinaldehyde dehydrogenase, expressed in the ventral region of the retina, Mech Dev. 98 (2000) 37–50.

[60] R.K. Kam, Y. Deng, Y. Chen, H. Zhao, Retinoic acid synthesis and functions in early embryonic development, Cell Biosci. 2 (2012) 11.

[61] N. Molotkova, A. Molotkov, G. Duester, Role of retinoic acid during forebrain development begins late when Raldh3 generates retinoic acid in the ventral subventricular zone, Dev Biol. 303 (2007) 601–610.

[62] A.M. Romani, Cellular magnesium homeostasis, Arch. Biochem. Biophys. 512 (2011) 1–23.

Table 1. Comparison of homologous residues lining the substrate-binding pocket of ALDH1A enzymes.

ALDH1A1	ALDH1A2	ALDH1A3
Leu114*	Pro131	Pro125
Asn121	Val138	Ile132
Gly125	**Gly142**	**Gly136**
Lys128	Lys145	Arg139
Thr129	**Thr146**	**Thr140**
Trp178	**Trp195**	**Trp189**
Ile304	Thr321	Thr315
Gly458	Asn475*	Asn469
Val459	Ala476*	Ala470
Val460	Leu477*	Leu471
Ser461	Asn478*	Tyr472
Ala462	**Ala479**	**Ala473**

*Asterisk indicates mutated residues. Bold face highlights strictly conserved residues.

Table 2. Kinetic constants of ALDH1A1, 1A2 and 1A3 with non-retinoid substrates.

Substrate and parameter	ALDH1A1	ALDH1A2	ALDH1A3
COFACTOR			
NAD^+			
K_m (μM)	25 ± 5	19 ± 2	130 ± 16
k_{cat} (min^{-1})	27 ± 1	262 ± 7	166 ± 7
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	1,050 ± 220	13,800 ± 1,300	1,300 ± 170
ALKANALS			
Butanal			
K_m (μM)	3.4 ± 0.7	19 ± 2	41 ± 4
k_{cat} (min^{-1})	23 ± 1	260 ± 6	75 ± 2
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	6,700 ± 1,400	13,700 ± 1,500	1,700 ± 170
Hexanal			
K_m (μM)	0.22 ± 0.05	3.0 ± 0.2	6.0 ± 0.3
k_{cat} (min^{-1})	31 ± 2	240 ± 35	154 ± 16
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	143,500 ± 31,000	82,000 ± 13,000	25,600 ± 3,000
Octanal			
K_m (μM)	3.0 ± 0.3	3.0 ± 0.8	85 ± 15
k_{cat} (min^{-1})	26 ± 1	212 ± 13	88 ± 5
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	8,700 ± 900	70,700 ± 19,300	1,035 ± 192
ALKENALS			
***Trans*-2-hexenal**			
K_m (μM)	0.56 ± 0.12	16 ± 2	30 ± 4
k_{cat} (min^{-1})	49 ± 3	57 ± 2	31 ± 1
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	87,200 ± 19,000	3,600 ± 500	1,024 ± 148
***Trans*-2-octenal**			
K_m (μM)	4.0 ± 0.5	8.0 ± 2.0	41 ± 5
k_{cat} (min^{-1})	7.0 ± 0.2	21 ± 1	6.0 ± 0.2
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	1,900 ± 260	2,600 ± 670	150 ± 20
4-Hydroxy-2-hexenal			
K_m (μM)	13 ± 3	NS	458 ± 68
k_{cat} (min^{-1})	14 ± 1		7.4 ± 0.5
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	1,100 ± 300	134	16 ± 3
4-Hydroxy-2-nonenal			
K_m (μM)	27 ± 6	7.5 ± 0.7	40 ± 8
k_{cat} (min^{-1})	10 ± 1	31 ± 1	7.4 ± 0.7
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	380 ± 90	4,100 ± 400	185 ± 40
Citral			
K_m (μM)	0.085 ± 0.01	54 ± 10	13 ± 2
k_{cat} (min^{-1})	1.00 ± 0.02	84 ± 7	7.0 ± 0.3
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	12,400 ± 1,600	1,600 ± 320	570 ± 90
ESTERASE ACTIVITY			
4-Nitrophenyl acetate			
K_m (μM)	56 ± 17	30 ± 2	59 ± 8.4
k_{cat} (min^{-1})	40 ± 4.6	85 ± 2	26 ± 0.5
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	710 ± 240	2,800 ± 200	450 ± 60

Enzymatic activity was measured fluorimetrically at 25°C. ALDH1A1 and ALDH1A2 were assayed in 50 mM HEPES, 0.5 mM EDTA, pH 8.0, 0.5 mM DTT. ALDH1A3 assays were performed in 50 mM HEPES, pH 8.0, 30 mM $MgCl_2$, 5 mM DTT. NAD^+ concentration was 500 μM. To calculate k_{cat} values the *Mr* used for ALDH1A1 was 220 kDa, for ALDH1A2 was 226.8 kDa and for ALDH1A3 was 224 kDa. NAD^+ constants were determined using 250 μM hexanal as a substrate; NS, not saturated up to 375 μM 4-hydroxy-2-hexenal.The corresponding k_{cat}/K_m value of ALDH1A2 for 4-hydroxy-2-hexenal was determined from the slope of V/[E] *versus* [S] linear plot.

Table 3. Kinetic constants of ALDH1A1, ALDH1A2 and ALDH1A3 with retinaldehyde isomers.

Substrate and parameter	ALDH1A1	ALDH1A2	ALDH1A3
all-*trans*-retinaldehyde			
K_m (μM)	11 ± 2	4.0 ± 1.0	2.4 ± 0.5
k_{cat} (min^{-1})	18 ± 2	20 ± 2	18 ± 1
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	1,600 ± 330	5,000 ± 1,300	7,600 ± 1,600
9-*cis*-retinaldehyde			
K_m (μM)	5.0 ± 1.5	1.2 ± 0.2	1.0 ± 0.1
k_{cat} (min^{-1})	20 ± 2	3.0 ± 0.1	0.090 ± 0.002
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	4,000 ± 1,400	2,500 ± 320	90 ± 13

Enzymatic activity was measured at 37°C by using the HPLC-based method. The assay buffer for each enzyme was the same as that for the fluorimetric assays.

Table 4. Kinetic constants of mutant ALDH1A1 and ALDH1A2.

Substrate and parameter	ALDH1A1	ALDH1A1 L114P	ALDH1A2	ALDH1A2 mutant[a]
NAD^+				
K_m (μM)	25 ± 5	74 ± 5	19 ± 2	8 ± 0.5
k_{cat} (min^{-1})	27 ± 1	28 ± 0.6	262 ± 7	470 ± 6
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	1,050 ± 220	380 ± 27	13,800 ± 1,300	56,000 ± 5000
Butanal				
K_m (μM)	3.4 ± 0.7	9.0 ± 1.0	19 ± 2	38 ± 2
k_{cat} (min^{-1})	23 ± 1	58 ± 2	260 ± 6	290 ± 3
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	6,700 ± 1400	6,300 ± 820	13,700 ± 1,500	7,500 ± 440
Hexanal				
K_m (μM)	0.22 ± 0.50	11 ± 1	3.0 ± 0.2	13 ± 1
k_{cat} (min^{-1})	31 ± 2	32 ± 1	240 ± 35	460 ± 6
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	143,500 ± 31,000	2,900 ± 190	82,000 ± 13,000	36,000 ± 1,700
Citral				
K_m (μM)	0.085 ± 0.010	8.0 ± 1.0	54 ± 10	1.0 ± 0.1
k_{cat} (min^{-1})	1.0 ± 0.1	10.0 ± 0.1	84 ± 7	18 ± 0.3
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	12,400 ± 1,600	1,200 ± 100	1,600 ± 320	16,000 ± 1,500
All-*trans*-retinaldehyde				
K_m (μM)	11 ± 2	20 ± 4	4.0 ± 1.0	5.0 ± 2.0
k_{cat} (min^{-1})	18 ± 2	19 ± 1.5	20 ± 2	53 ± 7
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	1,600 ± 330	940 ± 210	5,000 ± 1,300	10,600 ± 4000
9-*Cis*-retinaldehyde				
K_m (μM)	5 ± 1.5	5 ± 1.2	1.2 ± 0.2	3 ± 0.5
k_{cat} (min^{-1})	20 ± 2	20 ± 1	3 ± 0.1	3.5 ± 0.14
k_{cat}/K_m ($mM^{-1}\cdot min^{-1}$)	4,000 ± 1,400	4,000 ± 993	2,500 ± 320	1,200 ± 200

[a]This mutant included the following substitutions: N475G, A476V, L477V and N478S. Enzymatic activity was measured by the HPLC-based method for all-*trans*-retinaldehyde and fluorimetrically for alkanals and citral. NAD^+ constants were determined using 250 μM hexanal as a substrate. To calculate k_{cat} values, the *Mr* values used, for wild-type and mutant ALDH1A1, was 220,000 and, for wild-type and mutant ALDH1A2, was 226,800.

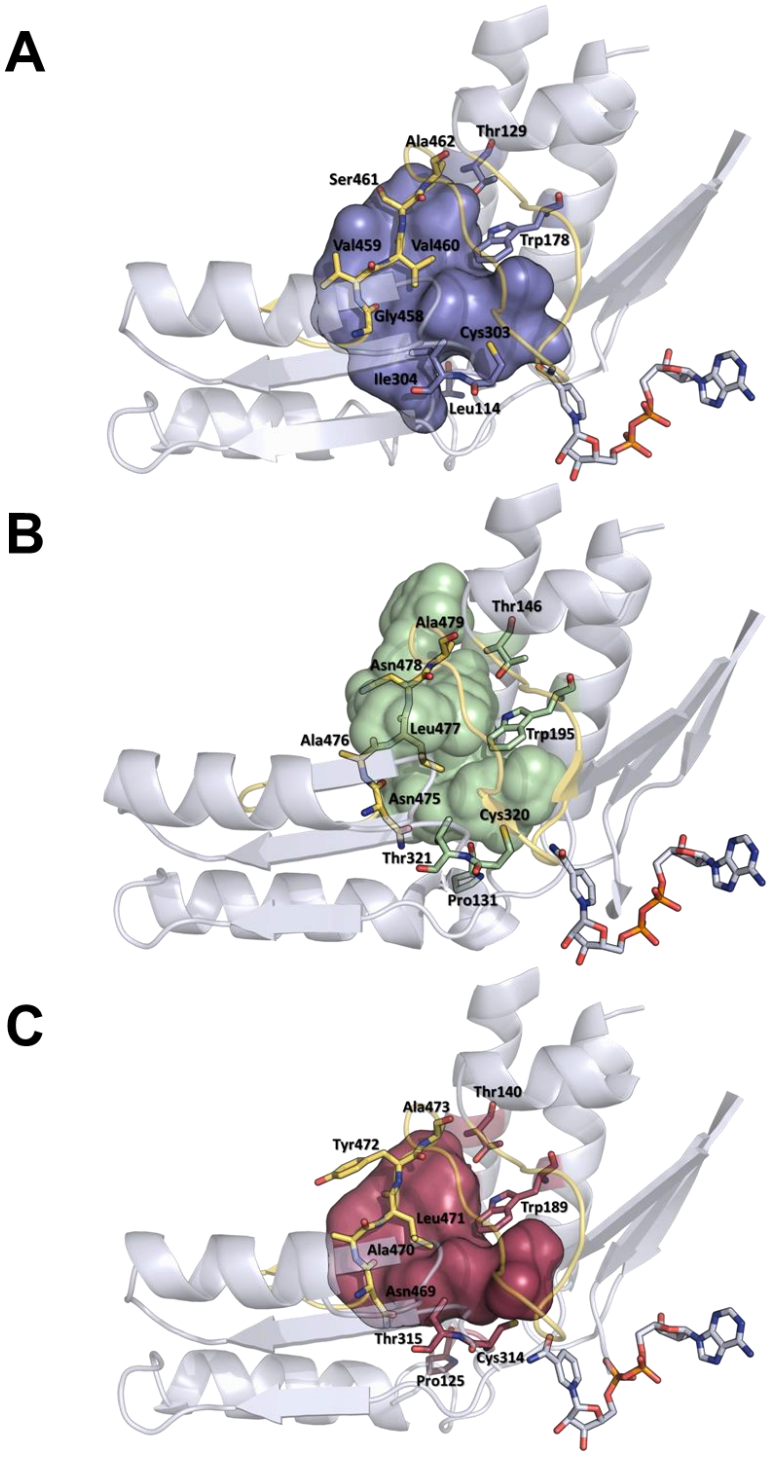

Figure 1. Crystallographic structures of the catalytic domain of the three human ALDH1A enzymes: **(A)** ALDH1A1, **(B)** ALDH1A and **(C)** ALDH1A3. The contour of the substrate-binding pockets of ALDH1A1, ALDH1A2 and ALDH1A3 is colored purple, green and burgundy red, respectively. The cofactor (NAD^+/NADH) is displayed in bluewhite. The residues lining the pocket, including the active-site cysteine, are indicated. The secondary structure is depicted in white ribbon representation while the 475-495 loop is shown in yellow.

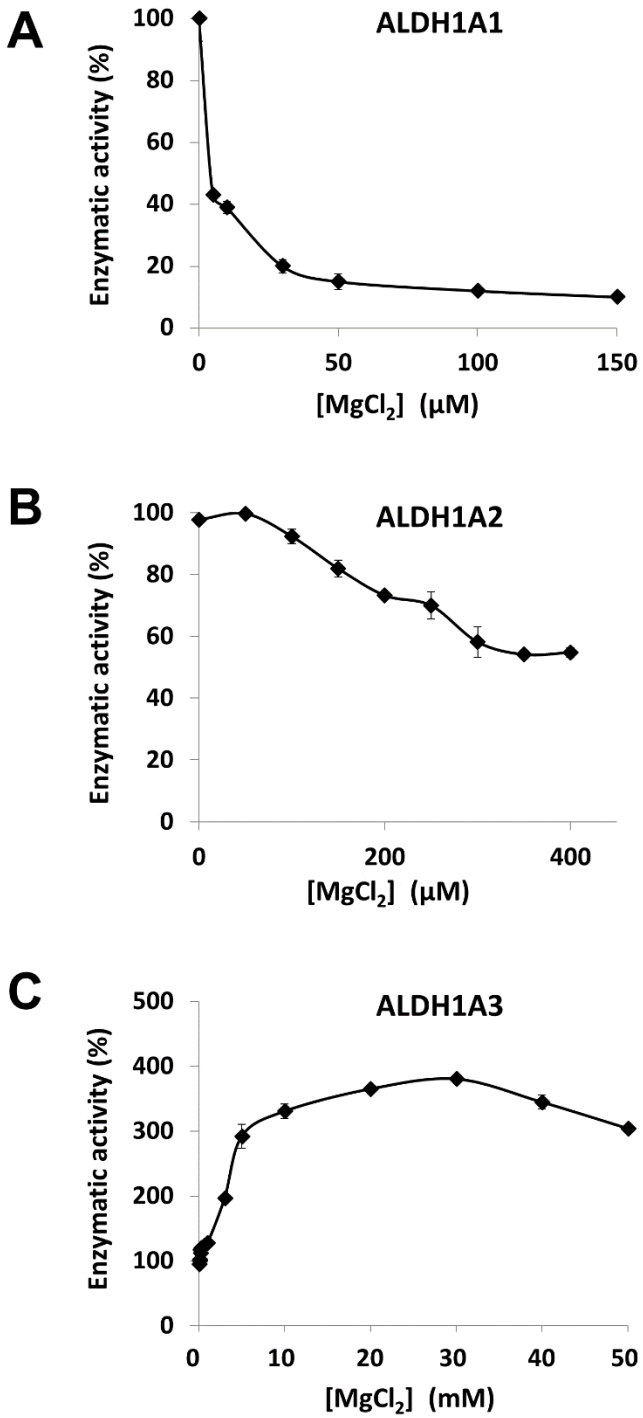

Figure 2. Effect of Mg^{2+} ions on the dehydrogenase activity of (A) ALDH1A1, (B) ALDH1A2 and (C) ALDH1A3. Enzymatic assays were carried out in 50 mM HEPES, 5 mM DTT, pH 8.0, with 0.5 mM NAD+ and 250 µM hexanal, at 25°C.

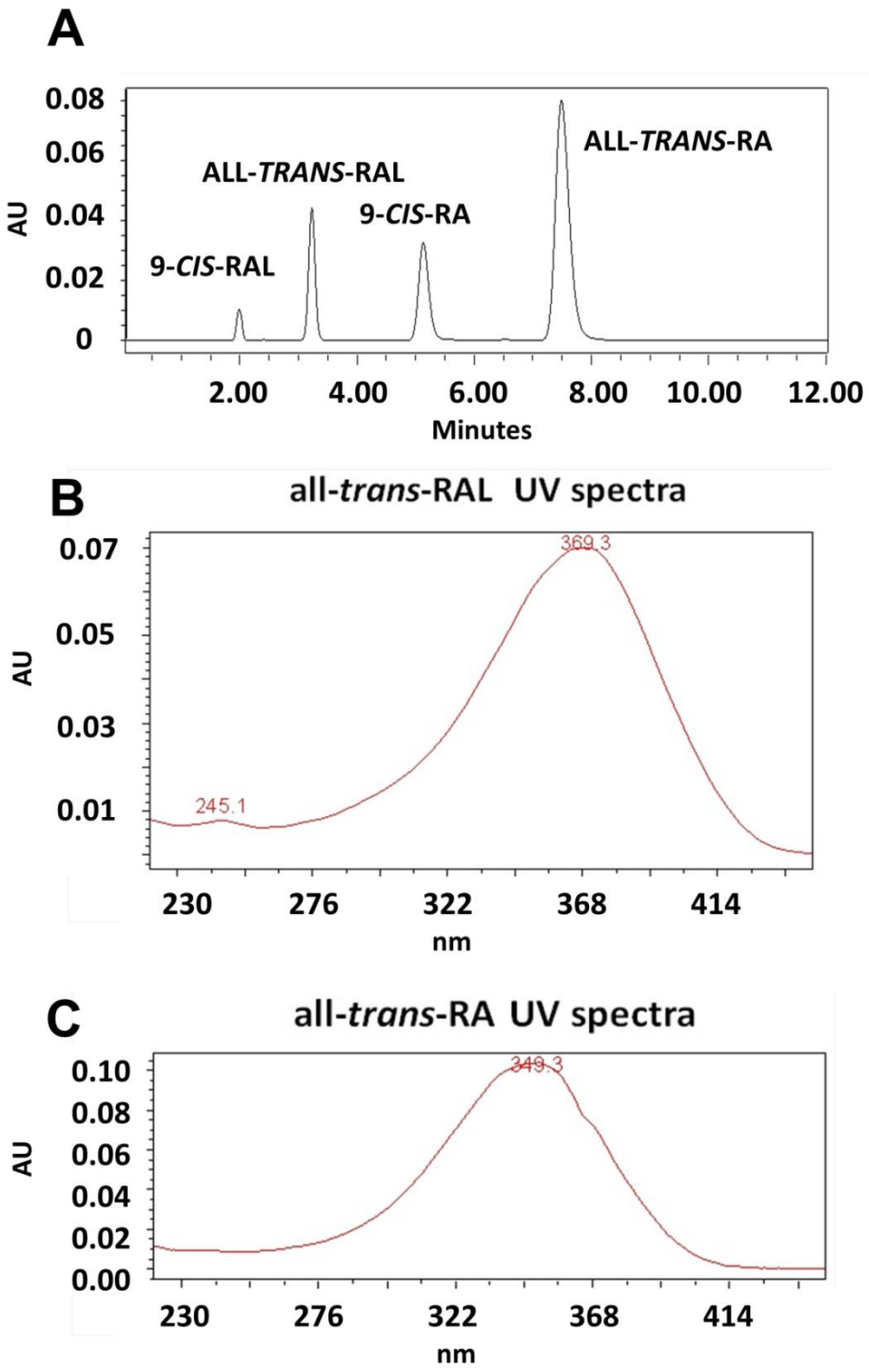

Figure 3. HPLC based method for ALDH activity determination with retinaldehyde. (A) Elution profile of all-*trans*- and 9-*cis*-retinoids and UV-vis spectra for (B) all-*trans*-retinaldehyde and (C) all-*trans*-retinoic acid. The HPLC chromatogram shows the separation of all-trans and 9-cis-retinaldehyde substrates and the corresponding RA products after the retinaldehyde dehydrogenase reaction catalyzed by ALDH1A1. Chromatogram recorded at 350 nm.

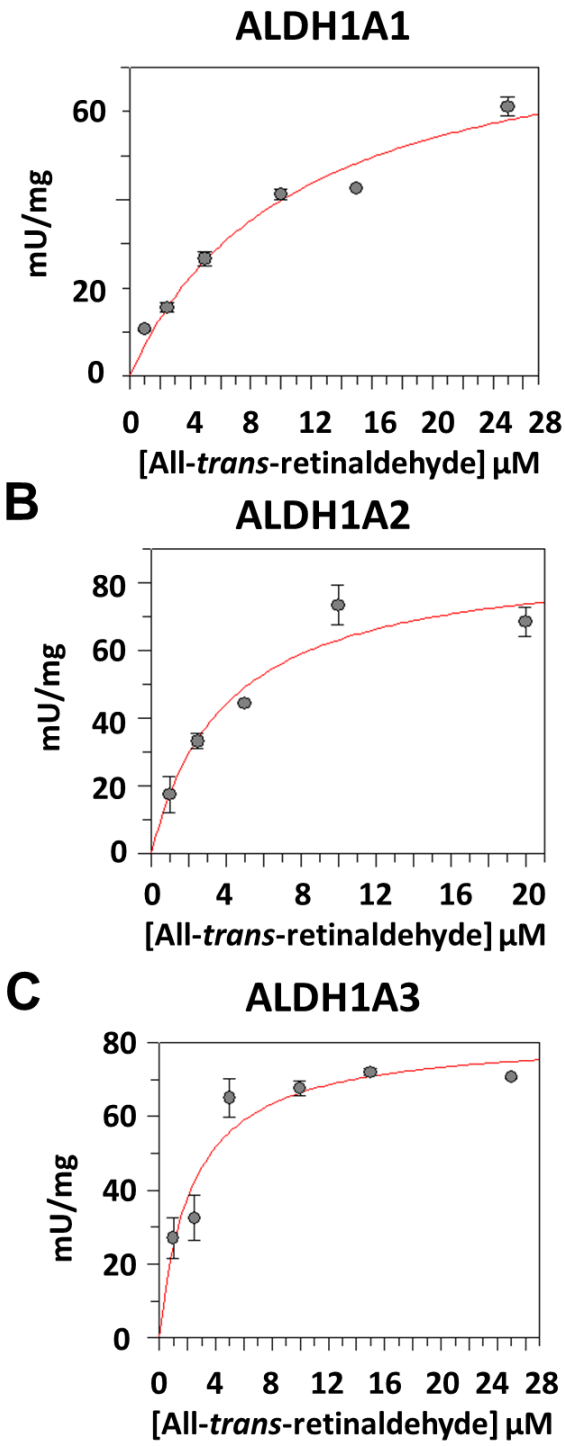

Figure 4. Michaelis-Menten plots for the kinetics of all-*trans*-retinaldehyde oxidation by (A) ALDH1A1, (B) ALDH1A2 and (C) ALDH1A3. The reaction was carried out in 50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0 (for ALDH1A1 and ALDH1A2) and 50 mM HEPES, 30 mM $MgCl_2$, 5 mM DTT, pH 8.0 (for ALDH1A3) at 37°C, and separation of retinoids by the HPLC-based method.

Table S1. Crystallographic structures of human ALDH1A enzymes deposited in the Protein Data Bank.

ALDH1A enzyme	cofactor	other ligands	resolution (Å)	PDB code	Released
1A1	−	−	1.74	4WJ9	Dec 2014
1A1	NADH	−	2.07	**4WB9**[a]	Dec 2014
1A1	−	CM026	1.80	4WP7	Feb 2015
1A1	−	CM053	1.95	4WPN	Feb 2015
1A1	NADH	CM037	1.85	4X4L	Feb 2015
1A1	−	BUC11	1.70	5L2M	Mar 2017
1A1	−	BUC22	2.05	5L2O	Mar 2017
1A1	−	BUC25	1.70	5L2N	Mar 2017
1A1	NADH	CM039	2.10	5TEI	Sept 2017
1A2	NAD^+	−	2.94	4X2Q	Dec 2015
1A2	NAD^+	WIN18,446	1.89	**6ALJ**[a]	Jan 2018
1A2	NAD^+	6-118	2.20	6B5G	Jan 2018
1A2	NAD^+	CM121	2.30	6B5H	Jan 2018
1A2	−	CM121	2.60	6B5I	Jan 2018
1A3	NAD^+	retinoic acid	2.90	**5FHZ**[a]	Nov 2016

[a]Structures used in this work shown in bold face. CM026: 8-{[4-(furan-2-ylcarbonyl)piperazin-1-yl]methyl}-1,3-dimethyl-7-(3-methylbutyl)-3,7-dihydro-1H-purine-2,6-dione; CM053: 1-{[1,3-dimethyl-7-(3-methylbutyl)-2,6-dioxo- 2,3,6,7-tetrahydro-1H-purin-8-yl]methyl}piperidine-4-carboxamide; CM037: ethyl ({4-oxo-3-[3-(pyrrolidin-1-yl)propyl]-3,4-dihydro[1]benzothieno[3,2-d]pyrimidin-2-yl}sulfanyl)acetate; BUC11: 2,3,5-trimethyl-6-[3-oxo-3-(piperidin-1-yl)propyl]-7H-furo[3,2-g][1]benzopyran-7-one; BUC 22: 7-(diethylamino)-4-methyl-2H-1-benzopyran-2-one; BUC25: 3-benzyl-4-methyl-2-oxo-2H-1-benzopyran-7-yl methanesulfonate; CM039: 6-{[(3-fluorophenyl)methyl]sulfanyl}-5-(2-methylphenyl)-2,5-dihydro-4H-pyrazolo[3,4-d]pyrimidin-4-one; WIN18,446: N,N'-(octane-1,8-diyl)bis(2,2-dichloroacetamide); 6-118: (3-ethoxythiophen-2-yl){4-[4-nitro-3-(pyrrolidin-1-yl)phenyl]piperazin-1-yl}methanone; CM121: 1-(4-cyanophenyl)-N-(3-fluorophenyl)-3-[4-(methylsulfonyl)phenyl]-1H-pyrazole-4-carboxamide.

Figure S1. Molecular structures of alkenal compounds used as ALDH1A substrates.