

IDENTIFICATION AND CHARACTERIZATION OF SMALL MOLECULE  
INHIBITORS OF ALDEHYDE DEHYDROGENASE 1A1

Cynthia A. Morgan

Submitted to the faculty of the University Graduate School  
in partial fulfillment of the requirements  
for the degree  
Doctor of Philosophy  
in the Department of Biochemistry and Molecular Biology,  
Indiana University

January 2015

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

---

Thomas D. Hurley, Ph.D., Chairman

---

Millie M. Georgiadis, Ph.D.

Doctoral Committee

---

Maureen A. Harrington, Ph.D.

December 11, 2014

---

William J. Sullivan, Ph.D.

## **Dedication**

To my parents for all their support and encouragement  
throughout the years

## Acknowledgements

First and foremost, I want to thank my advisor, Dr. Thomas Hurley, for being an outstanding mentor and friend. At the start of my rotation, he trusted me to conduct the expensive and challenging high throughput screen that became the foundation of this dissertation. Throughout my time in his laboratory, he continued to allow me to handle projects independently, building my confidence as a scientist-in-training, but he was always there when I needed help. I am sincerely grateful for his patience, enthusiasm, and kindness. I could not have asked for a better advisor for my graduate studies.

I would also like to thank my committee members for all their time, suggestions, and insight over the last four years. I thank Dr. Millie Georgiadis for constantly challenging me in X-ray crystallography. I thank Dr. Maureen Harrington and Dr. William Sullivan for forcing me to think outside the *in vitro* world and to address the ramifications of my project within the organism, from cell culture to polar bears.

Science is a highly collaborative effort and I would like to thank the following individuals who contributed to this project. The Indiana University Chemical Genomics Core Facility, in particular Dr. Lan Chen and Andrea Gunawan, for providing access to their chemical libraries and equipment for the high throughput screen. I want to thank Dr. Tsuyoshi Imasaki for assistance in collecting data for the ALDH1A1 apo structure, and Dr. Karl Dria for assistance with LC/MS. I thank Dr. Harrington for allowing me access to her cell culture facility. As my own endeavors in cell culture were unsuccessful, a special thanks to Dr. Daniela Matei and Dr. Salvatore Condello at Indiana University School of Medicine and Dr. Ronald Buckanovich at the University of Michigan School of Medicine for their expertise in testing our compounds in ovarian cancer cells.

I want to thank my colleagues in the South Wing of the Biochemistry Department. Dr. Bibek Parajuli for assisting me during my rotation and helping me get started in the Hurley lab. Dr. Vimbai Chikwana and Dr. May Khanna for always having an answer to my numerous questions. Dr. Ann Kimble-Hill, Cameron Buchman, Krishna Kishore Mahalingan, and Lanmin Zhai, along with Dr. Kentaro Yamada, Qiuqia Chen, and Isha

Singh for your friendship and support, as well as providing humor and a much needed distraction from a busy day of science.

I also want to thank Dr. Debomoy Lahiri, Dr. Jay Simon, Dr. Martin Bard, Dr. John Luke, Dr. Richard Jaskunas, Dr. Clark Wells, and Col. Michael Yard for their mentoring and career guidance over the years.

Special thanks to the Department of Biochemistry and Molecular Biology front office, Sandy, Patty, Sheila, Darlene, Melissa, and Jack, for their assistance and constant supply of coffee and cookies.

This research was supported by the U.S. National Institute of Health [R01AA018123] to TDH. Results shown in this report are derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

Cynthia A. Morgan

IDENTIFICATION AND CHARACTERIZATION OF SMALL MOLECULE  
INHIBITORS OF ALDEHYDE DEHYDROGENASE 1A1

The human genome encodes 19 members of the aldehyde dehydrogenase (ALDH) superfamily, critical enzymes involved in the metabolism of aldehyde substrates. A major function of the ALDH1A subfamily is the oxidation of retinaldehyde to retinoic acid, a key regulator of numerous cell growth and differentiation pathways. ALDH1A1 has been identified as a biomarker for both normal stem cells and cancer stem cells. Small molecule probes are needed to better understand the role of this enzyme in both normal and disease states. However, there are no commercially available, small molecules that selectively inhibit ALDH1A1. Our goal is to identify and characterize small molecule inhibitors of ALDH1A1 as chemical tools and as potential therapeutics. To better understand the basis for selective inhibition of ALDH1A1, we characterized N,N-diethylaminobenzaldehyde (DEAB) which is a commonly used inhibitor of ALDH1A1 and purported to be selective. DEAB serves as the negative control for the Aldefluor assay widely utilized to identify stem cells. Rather than being a selective inhibitor for ALDH1A1, we found that DEAB is a slow substrate for multiple ALDH isoenzymes, and depending on the rate of turnover, DEAB behaves as either a traditional substrate or as an inhibitor. Due to its very slow turnover, DEAB is a potent inhibitor of ALDH1A1 with respect to propionaldehyde oxidation, but it is not a good candidate for the development of selective ALDH1A1 inhibitors because of its promiscuity. Next, to discover novel

selective inhibitors, we used an *in vitro*, high-throughput screen of 64,000 compounds to identify 256 hits that either activate or inhibit ALDH1A1 activity. We have characterized two structural classes of compounds, CM026 and CM037, using enzyme kinetics and X-ray crystallographic structural data. Both classes contained potent and selective inhibitors for ALDH1A1. Structural studies of ALDH1A1 with CM026 showed that CM026 binds at the active site, and its selectivity is achieved by a single residue substitution. Importantly, CM037 selectively inhibits proliferation of ALDH+ ovarian cancer cells. The discovery of these two selective classes of ALDH1A1 inhibitors may be useful in delineating the role of ALDH1A1 in biological processes and may seed the development of new chemotherapeutic agents.

Thomas D. Hurley, Ph.D., Chairman

## Table of Contents

List of Tables .....	x
List of Figures .....	xi
I. Introduction .....	1
A. Aldehydes .....	1
B. The Aldehyde Dehydrogenase Superfamily .....	5
C. Aldehyde Dehydrogenase 1A1 .....	12
D. ALDH1A1 and Parkinson's disease .....	14
E. ALDH1A1 and Cancer .....	16
F. ALDH1A1 and Obesity .....	19
G. The Aldefluor Assay .....	20
H. ALDH1A1 Activators and Inhibitors .....	22
I. Hypothesis and Approach .....	26
II. Materials and Methods .....	30
A. Materials .....	30
B. Methods .....	30
1. Production of ALDH Isoenzymes .....	30
2. Generation of wild-type ALDH1A1 and the G458N mutant .....	33
3. X-ray crystallography .....	33
4. Esterase High-Throughput Screen .....	35
5. Aldehyde Oxidation Activity Assays for ALDH Isoenzymes .....	38
6. Steady-State Enzyme Kinetics .....	39
7. Wavelength Scans to Monitor DEAB Oxidation .....	42
8. Mass Spectrometry .....	42
III. Characterization of Diethylaminobenzaldehyde .....	43
A. Results .....	43
1. DEAB as a Substrate for ALDH Isoenzymes .....	43
2. DEAB as an Inhibitor of ALDH Isoenzymes .....	46
3. Mass Spectrometry of ALDH-DEAB Complexes .....	51
B. Discussion .....	55



IV. Discovery of Novel, Selective Inhibitors of ALDH1A1 .....	60
A. Results .....	60
1. X-ray Crystallography of apo-ALDH1A1 and ALDH1A1-NADH Complex .....	60
2. Esterase High-Throughput Screen.....	64
3. Selectivity of Compounds for ALDH1A1 Aldehyde Oxidation .....	66
4. Steady-State Kinetics Optimization.....	71
5. Characterization of the CM037 Class of Compounds .....	74
6. Characterization of the CM026 Class of Compounds .....	78
7. X-ray Crystallography of ALDH1A1 with CM026, CM053, and CM037 .....	82
8. Characterization of ALDH1A1 G458N Mutant .....	87
B. Discussion .....	88
V. Conclusion and Future Directions.....	104
Supplementary Table 1: Activators of ALDH1A1 esterase activity. ....	110
Supplementary Table 2: Inhibitors of ALDH1A1 esterase activity.....	121
References .....	122
Curriculum vitae	

## List of Tables

Table 1: Common sources of exogenously and endogenously produced aldehydes.....	2
Table 2: The 19 Human Aldehyde Dehydrogenase Isoenzymes .....	9
Table 3: Solutions used in the preparation of ALDH1A1 protein.....	31
Table 4: Kinetic parameters for ALDH3A1 oxidation of DEAB .....	45
Table 5: Effect of DEAB on ALDH isoenzymes.....	48
Table 6: Data collection and refinement statistics.....	62
Table 7: IC <sub>50</sub> values and structures of the top 19 compounds that inhibit ALDH1A1....	72
Table 8: SAR for CM026 Analogs. ....	81
Table 9: Data collection and refinement statistics.....	82
Table 10: Kinetic parameters of ALDH1A1 WT and mutant G458N.....	88
Table 11: Structural classes of hit compounds. ....	93

## List of Figures

Figure 1: Aldehyde .....	1
Figure 2: Protein carbonylation. ....	3
Figure 3: Aldehyde-induced DNA damage. ....	4
Figure 4: Aldehyde detoxication enzymes.....	5
Figure 5: ALDH Nomenclature .....	6
Figure 6: Mechanisms of aldehyde oxidation and ester hydrolysis of ALDH. ....	8
Figure 7: Ribbon diagram of the ALDH2 monomer and homotetramers.....	11
Figure 8: ALDH in the retinoic acid signaling pathway.....	13
Figure 9: Dopamine synthesis and inactivation. ....	15
Figure 10: The Aldefluor Assay. ....	21
Figure 11: Reactions used to discover ALDH1A1 modulators. ....	25
Figure 12: Comparison of the active sites of three ALDH isoenzymes. ....	27
Figure 13: Small molecule discovery. ....	28
Figure 14: ALDH1A1 protein purification. ....	32
Figure 15: Esterase high-throughput screen. ....	35
Figure 16: Esterase HTS plate set-up.....	37
Figure 17: Aldehyde oxidation activity assays. ....	38
Figure 18: Absorption characteristics of DEAB and diethylaminobenzoic acid.....	43
Figure 19: Oxidation of DEAB in the absence of enzyme. ....	44
Figure 20: Wavelength scans of ALDH isoenzymes with DEAB.....	45
Figure 21: Substrate saturation curve for ALDH3A1 with DEAB as a substrate. ....	46
Figure 22: Selectivity of DEAB with eight ALDH isoenzymes.....	47
Figure 23: Characterization of ALDH1A1 with DEAB. ....	47
Figure 24: ALDH2 and DEAB. ....	48
Figure 25: Time-dependent inhibition of human ALDH1A2 with DEAB. ....	49
Figure 26: Time-dependent inhibition of human ALDH2 with DEAB.....	50
Figure 27: Modification of human ALDH2 with DEAB.....	51
Figure 28: Modification of human ALDH1A1 with DEAB.....	52
Figure 29: Modification of human ALDH1A1 with benzaldehyde.....	53
Figure 30: Modification of human ALDH1A2 with DEAB.....	54

Figure 31: Effects of dialysis on DEAB-induced inhibition.....	55
Figure 32: Proposed mechanism of action of DEAB on ALDH enzymes. ....	56
Figure 33: Evolution of ALDH1A1 crystallization conditions.....	60
Figure 34: Structure of human ALDH1A1 (N121S) apo-enzyme.....	63
Figure 35: Structure of ALDH1A1 (WT) with reduced cofactor NADH.....	63
Figure 36: Comparison of the cofactor binding sites of ALDH1A1 and ALDH2.....	64
Figure 37: Z-factor determination for HTS esterase screen.....	65
Figure 38: Representative plate from HTS. ....	66
Figure 39: Effect on aldehyde oxidation of hit compounds from HTS. ....	68
Figure 40: IC <sub>50</sub> curves and structures of the four most potent ALDH1A1 inhibitors.....	70
Figure 41: Calculation of K <sub>m</sub> 's for steady-state kinetics with ALDH1A1. ....	73
Figure 42: Characterization of CM037, a selective inhibitor of ALDH1A1. ....	74
Figure 43: Basic structure of CM037 Analogs. ....	75
Figure 44: IC <sub>50</sub> Curves for CM037 and two of its analogs.....	76
Figure 45: CM037 and its analogs.....	77
Figure 46: Structure of CM026.....	78
Figure 47: Selectivity and potency of CM026.....	78
Figure 48: Mode of Inhibition for CM026, a selective inhibitor of ALDH1A1.....	79
Figure 49: General structure of CM026 analogs.....	79
Figure 50: Common structural analogs of CM026. ....	80
Figure 51: Structure of ALDH1A1 with CM026.....	83
Figure 52: Structure of human ALDH1A1 with CM053.....	84
Figure 53: Structure of ALDH1A1 N121S with CM037.....	85
Figure 54: Structural basis of selectivity of CM026 for ALDH1A1. ....	86
Figure 55: Selectivity of compounds for WT vs G458N mutant.....	87
Figure 56: Surface topography of the cofactor binding site for 3 ALDH isoenzymes. ....	88
Figure 57: Esterase activators as inhibitors of aldehyde oxidation.....	91
Figure 58: Schematic representation of CM026 inhibition of ALDH1A1. ....	94
Figure 59: Schematic representation of CM037 inhibition of ALDH1A1 .....	96
Figure 60: Possible hydrolysis of CM037 by ALDHs.....	97
Figure 61: Comparison of the active site topography of human ALDHs. ....	98

Figure 62: Comparison of compound binding to ALDH isoenzymes. ....	99
Figure 63: The effect of CM037 in ovarian cancer cells. ....	101
Figure 64: Effect of various compounds on the proliferation of OC cell. ....	102

## I. Introduction

### A. Aldehydes

Aldehydes are organic compounds with the general formula RCHO, with CHO representing a terminal carbonyl group and the R group ranging from a hydrogen atom to complex organic compounds (Figure 1). They are found extensively in the environment and are products of numerous biological pathways. Although small reserves of aldehydes exist

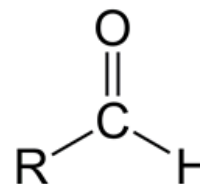


Figure 1: Aldehyde

naturally, such as the production of formaldehyde from the photochemical oxidation of methane in the atmosphere, most exogenous sources of aldehydes have an anthropogenic origin. Motor vehicle exhaust is a leading contributor, releasing formaldehyde, acetaldehyde, and acrolein as well as other carbonyls, plus hydrocarbons that can undergo photochemical oxidation to aldehydes. Aldehydes are used or released in a variety of industrial applications, including the production of plastics and resins, and in agriculture as both a preservative in animal feed and fumigant for pest control. The indoor environment is estimated to have a 4- to 10-fold higher level of aldehydes than outdoor air, with cigarette smoke, furniture, paints, cleaning agents, and cooking fumes all contributing to the aldehyde load<sup>1</sup>. Aldehydes, such as citral, acetaldehyde, and benzaldehyde, are found naturally or added to foods and beverages. A number of drugs are metabolized via an aldehyde intermediate, including ethanol, the cancer drug cyclophosphamide<sup>2,3</sup>, the sedative chloral hydrate<sup>4</sup>, and the reverse transcriptase inhibitor Abacavir<sup>5</sup>. Endogenously, lipid peroxidation, the oxidative degradation of membrane lipids, is estimated to produce over 200 aldehydes, including 4-hydroxynonenal (4-HNE) and malondialdehyde (MDA)<sup>6</sup>. Aldehydes are also produced during both amino acid and carbohydrate metabolism, generating glyceraldehyde, glycolaldehyde, methylglyoxal, and glutamic- $\gamma$ -semialdehyde. Common sources of aldehydes are summarized in Table 1.

Aldehydes play vital roles in normal biological processes. The retinaldehyde 11-cis-retinal binds the protein opsin, forming the photoreceptor rhodopsin that is critical for animal vision<sup>7</sup>. Retinaldehydes are also the precursors to retinoic acids, with both substrate and product serving as ligands of the retinoic acid and/or retinoic-X receptors

Table 1: Common sources of exogenously and endogenously produced aldehydes.

<b>Exogenous Source</b>	<b>Aldehyde</b>
Foods & Beverages	Benzaldehyde, crotonaldehyde, acrolein, acetaldehyde
Combustion	Formaldehyde, acetaldehyde, acrolein
Cigarettes	$\gamma$ -3-Pyridyl- $\gamma$ -methylaminobutyraldehyde, formaldehyde
<b>Endogenous Source</b>	<b>Aldehyde</b>
Lipid peroxidation	4-hydroxynonenal, malondialdehyde, hexanal
Vitamin A metabolism	Retinaldehyde
Dopamine catabolism	3,4-Dihydroxyphenylacetaldehyde
Corticosteroid catabolism	21-Dehydrocorticosteroids
Proline biosynthesis	Glutamic- $\gamma$ -semialdehyde
Sorbitol metabolism	Glycolaldehyde, glyceraldehyde, glyoxal
Putrescine catabolism	$\gamma$ -amino butyraldehyde

(RAR and RXR), transcription factors that are critical for cellular growth and differentiation pathways<sup>8</sup>. Glyceraldehyde-3-phosphate is an intermediate in glycolysis and gluconeogenesis, as well as in photosynthesis, serving as a vital component of cellular energy homeostasis<sup>7</sup>.

Although aldehydes are critical for biological functions, the electrophilic carbonyl group is highly reactive, and aldehydes are capable of forming adducts with various cellular components including proteins, nucleic acids, and glutathione (GSH), leading to dysfunction and cell death. Unlike reactive oxygen and nitrogen species, aldehydes have relatively long lifespans, and they can be transported or diffuse to sites far removed from where they are produced. Protein carbonylation is the formation of a covalent adduct between an aldehyde and a protein involving either a Schiff base formation via the  $\epsilon$ -amino of lysine or a Michael addition to cysteine, histidine, or lysine (Figure 2)<sup>9,10</sup>. Moderate carbonylation can be processed by the cell, but as the load of carbonylated proteins increases, they can aggregate and resist proteolysis, leading to protein dysfunction, cell death, and ultimately contribute to disease progression<sup>11</sup>. Protein carbonylation is elevated during oxidative stress, a process that generates numerous toxic aldehydes, and is associated with many diseases including rheumatoid arthritis, cardiovascular disease, Parkinson's disease, and Alzheimer's disease. Aldehydes and

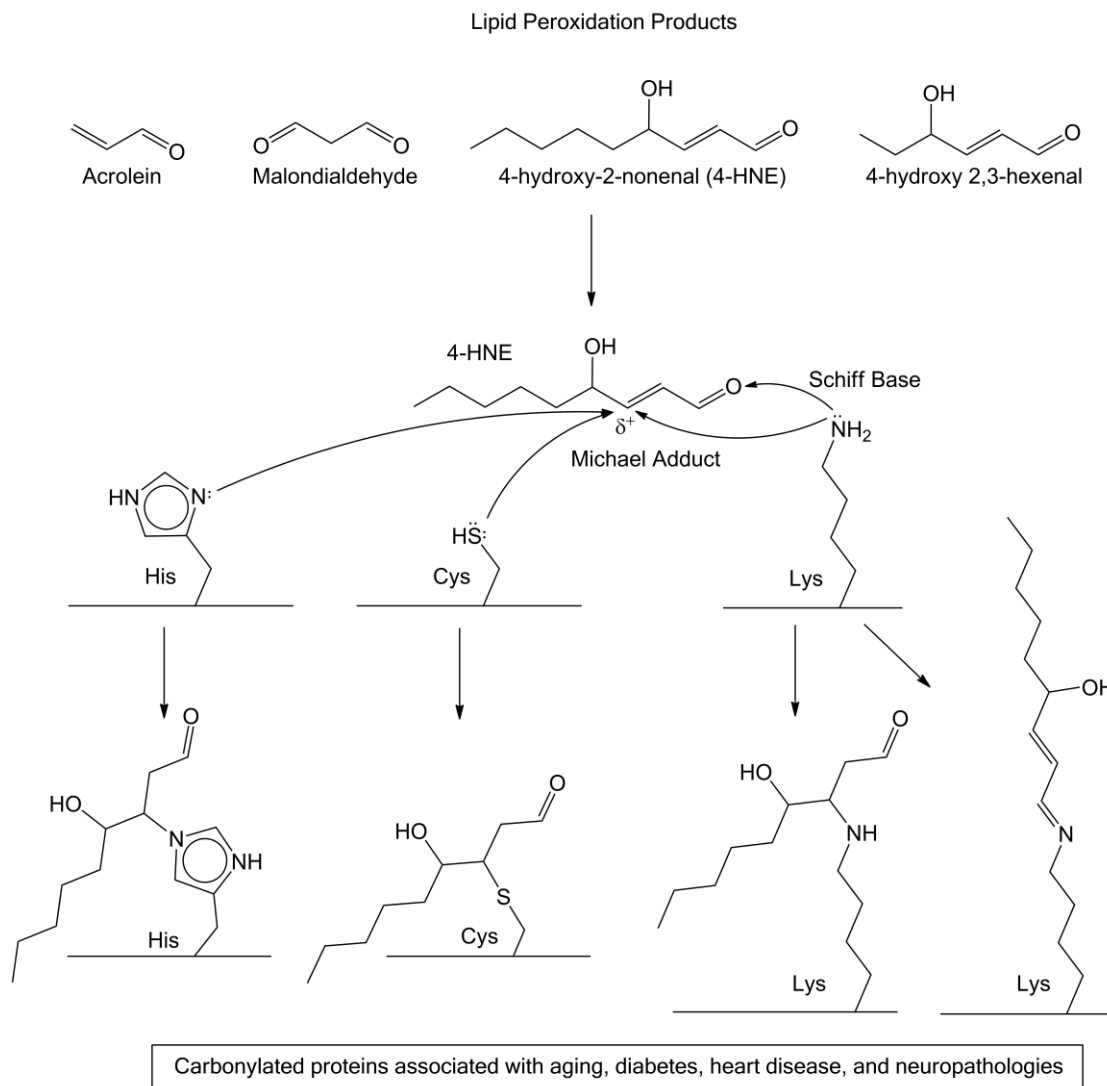


Figure 2: Protein carbonylation.

Formation of a covalent adduct between an aldehyde and a protein. Lipid peroxidation produces numerous aldehydes, including 4-HNE, which is capable of forming adducts, most often on histidine, lysine, and cysteine residue of the protein (modified from Curtis *et al*<sup>9</sup> and Ross *et al*<sup>10</sup>).

other carbonyls have been shown to inactivate glutathione reductase, glutathione peroxidase, and glutathione transferases, increasing cellular oxidative stress<sup>1</sup>. Aldehydes can also form adducts with DNA, leading to DNA-DNA and DNA-protein crosslinks, mutagenesis, chromosomal aberrations and other DNA damage (Figure 3)<sup>12</sup>. Cellular defense mechanisms exist to minimize damage, and mutations in these repair pathways have been linked to increased cancer rates<sup>12,13</sup>.



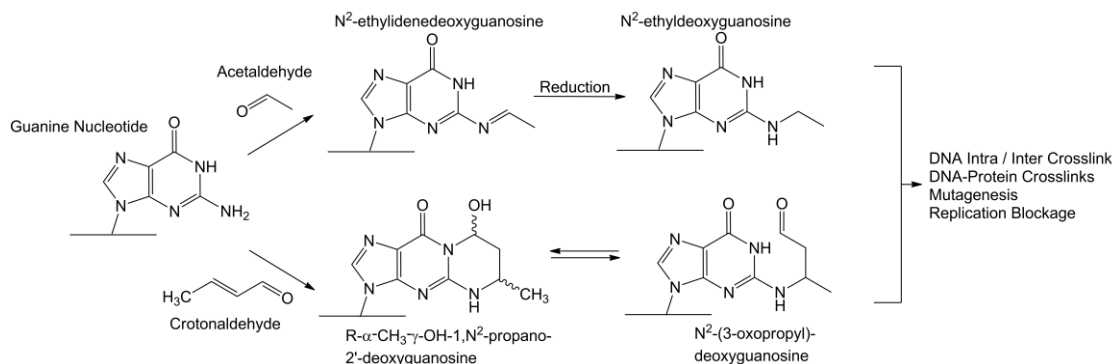
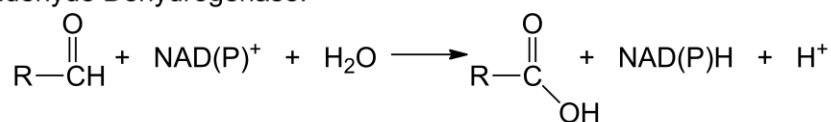


Figure 3: Aldehyde-induced DNA damage (modified from Brooks *et al*<sup>12</sup>).

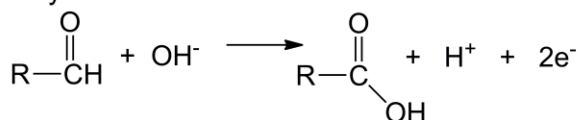
Due to the cytotoxic nature of aldehydes, their detoxification is critical for cellular homeostasis. To decrease the aldehyde burden in a cell, aldehydes are reduced to an alcohol, oxidized to an acid, or eliminated via the antioxidant glutathione. The main aldehyde reduction enzyme systems are alcohol dehydrogenases and aldo-keto reductases, while the main aldehyde oxidation enzyme systems are aldehyde/xanthine oxidase and aldehyde dehydrogenase (Figure 4). Alcohol dehydrogenases catalyze the reversible reduction of aldehydes and ketones to an alcohol, with the direction heavily dependent on the  $\text{NAD}^+$  to  $\text{NADH}$  ratio; given the 500:1 ratio of  $\text{NAD}^+$  to  $\text{NADH}$  commonly found in cells, the oxidation of an alcohol to a carbonyl is the dominant reaction<sup>14,15</sup>. The aldo-keto reductase superfamily also catalyzes the reversible reduction of a carbonyl to its corresponding alcohol, and this superfamily plays a large role in the removal of reactive aldehydes and other carbonyl-containing organic compounds<sup>16</sup>. Aldehyde and xanthine oxidase metabolize a variety of aromatic and heterocyclic aldehydes to their corresponding carboxylic acid<sup>17</sup>. The aldehyde dehydrogenase superfamily catalyzes the irreversible,  $\text{NAD(P)}^+$ -dependent oxidation of an aldehyde to its corresponding carboxylic acid or ester  $\text{CoA}$ <sup>18</sup>. Finally, reactive aldehydes are also eliminated via conjugation to the tripeptide glutathione<sup>19</sup>. Although all these pathways contribute to reducing the cellular aldehyde burden, the aldehyde dehydrogenase superfamily is capable of oxidizing a wide-variety of aldehyde substrates, both aliphatic and aromatic, and plays a critical role in both the cellular protection from reactive aldehydes and the generation of biologically important compounds.

### Aldehyde Oxidation

Aldehyde Dehydrogenase:

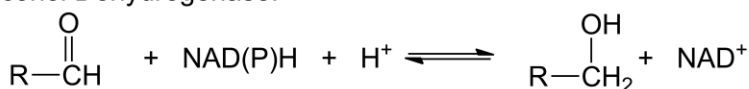


Aldehyde / Xanthine Oxidase:



### Aldehyde Reduction

Alcohol Dehydrogenase:



Aldo-Keto Reductase

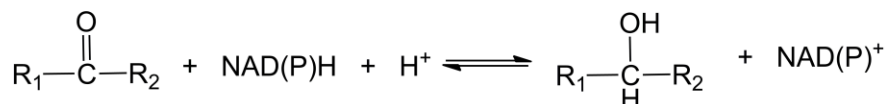


Figure 4: Aldehyde detoxication enzymes.

## **B. The Aldehyde Dehydrogenase Superfamily**

Members of the aldehyde dehydrogenases (ALDH) superfamily of enzymes are found in all three taxonomic domains, *Archaea*, *Eubacteria*, and *Eukarya*, indicating that this superfamily of proteins plays a critical evolutionary role<sup>20</sup>. The Ninth International Symposium on Enzymology and Molecular Biology of Carbonyl Metabolism established a standardized gene nomenclature system for the ALDH superfamily based on divergent evolution and amino acid identity (Figure 5)<sup>21</sup>. Members of the same ALDH family have more than 40% protein sequence identity, while members of the same subfamily have more than 60% protein sequence identity. A nomenclature exception has been made for ALDH2, an enzyme involved in ethanol metabolism with an immense presence in the literature record. ALDH2's longstanding name has been grandfathered in despite its amino acid sequence placing it in the ALDH1A subfamily. Currently the eukaryotic ALDH superfamily has 24 families, with 11 families represented in the human genome<sup>22</sup>. The human genome contains at least 19 ALDH genes, most of which catalyze the

NAD(P)<sup>+</sup>-dependent oxidation of a wide variety of endogenous and exogenous aldehydes to their carboxylic acids. ALDHs differ in their tissue distribution, subcellular location, structure, and substrate preference. They are found in numerous subcellular locations, including the cytoplasm, mitochondria, endoplasmic reticulum, and the nucleus. ALDHs are critical enzymes in biological functions, as well as in the defense

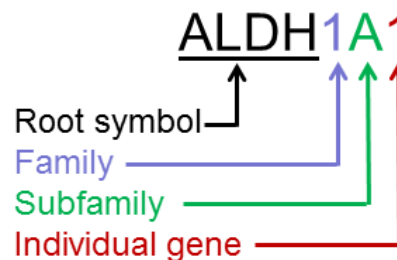


Figure 5: ALDH Nomenclature  
Members of the same family have >40% protein sequence identity; members of the same subfamily share > 60% identity.

against toxic aldehydes<sup>23</sup>. Through their role in aldehyde oxidation, ALDHs are involved in the synthesis of a number of key molecules, including retinoic acid, betaine, and tetrahydrofolate, as well as the removal of toxic aldehydes. A second enzymatic role for ALDHs is the hydrolysis of esters, although the physiological importance of its esterase activity is not well understood<sup>24</sup>. Some members of the ALDH superfamily have functions removed from an enzymatic role. ALDHs in ocular tissue act as crystallins, structural proteins capable of protecting the eye by absorbing ultraviolet radiation, while others serve as binding proteins for hormones and other small molecules<sup>6</sup>. A summary of the 19 human ALDH proteins is shown in Table 2.

Due to their involvement in a variety of physiological processes, mutations in human ALDH genes have been linked to a number of diseases (Table 2). The ALDH1A subfamily of proteins, via their production of retinoic acid, is critical during embryogenesis, and mutations in this subfamily have been associated with congenital heart disease and spina bifida<sup>6,25</sup>. As crystallin proteins, ALDH1A1 and ALDH3A1 protect ocular tissue from reactive oxygen species and absorb damaging UV radiation<sup>26,27</sup>. ALDH1A1<sup>-/-</sup> and ALDH3A1<sup>-/-</sup> single knockout mice as well as the ALDH1A1<sup>-/-</sup> / ALDH3A1<sup>-/-</sup> double knockout mice had premature cataract formation<sup>28</sup>. ALDH2 plays a major role in the detoxication of aldehydes in response to ischemia and of acetaldehyde during ethanol metabolism<sup>6</sup>. Mutations in the ALDH2 gene can lead to alcohol flushing syndrome following alcohol consumption due to elevated blood

acetaldehyde levels<sup>29</sup>. ALDH1A1 and ALDH2 are both involved in dopamine metabolism, a neurotransmitter linked to Parkinson's disease<sup>30</sup> and drugs of addiction<sup>31</sup>. It is common for mutations in ALDH genes to result in neurological defects. Mutations in the ALDH5A1 gene, the last step in the metabolism of the neurotransmitter GABA, are responsible for the autosomal recessive disorder  $\gamma$ -hydroxybutyric aciduria, marked by neurological and cognitive defects<sup>32</sup>. ALDH3A2, is involved in the oxidation of fatty alcohols to fatty acids, and mutations in this gene are associated with Sjögren-Larsson syndrome, an autosomal recessive disorder that causes skin, eye, and neurological problems<sup>33</sup>. ALDH4A1 is involved in proline degradation and mutations cause type II hyperprolinemia, marked by mental retardation and seizures<sup>34</sup>. ALDH18A1 is important for proline and arginine synthesis, with mutations resulting in low levels of these critical amino acids, characterized by metabolic and neurological disorders, including hypoprolinemia<sup>35</sup>. Several ALDH isoenzymes (ALDH1A1, ALDH1A2, ALDH1A3, ALDH1L1, ALDH2, ALDH3A1, ALDH4A1, and ALDH7A1) have been associated with cancer, and over-expression of ALDHs in cancer cells is linked to poor prognosis and an increased resistance to certain cancer drugs, including cyclophosphamide, that utilize ALDHs for drug metabolism<sup>3,6</sup>. In mammals and fish, ALDH16A1 is unique as the active site lacks key catalytic residues, rendering this enzyme inactive with respect to aldehyde oxidation; however, ALDH16A1 may function as a binding protein and has been linked to gout<sup>36</sup>.

The primary function of almost all ALDHs is the oxidization of a wide range of aldehydes to their corresponding carboxylic acid or to CoA esters utilizing the cofactor NAD(P)<sup>+</sup>. However, early sequence alignment of 145 ALDHs indicated only a limited number of residues are conserved in catalytically active enzymes, including the catalytic cysteine and residues involved in cofactor binding<sup>37</sup>. Using the mature ALDH2 sequence as reference, Cys302 is the active site nucleophile<sup>38</sup> while Glu268 acts as a general base<sup>39</sup>. The basic mechanism of the reaction catalyzed by ALDHs is shown in Figure 6A. The aldehyde oxidation mechanism starts with binding of the cofactor NAD(P)<sup>+</sup> and concludes with release of the reduced cofactor, but there are differences in the rate-limiting step among the different ALDH enzymes. For example, the rate-limiting step for

the ALDH1A subfamily is cofactor dissociation, deacylation for ALDH2, and hydride transfer for ALDH3A1<sup>40-42</sup>. Many ALDHs also possess an NAD(P)<sup>+</sup>-independent esterase activity<sup>24</sup>, and this esterase activity is proposed to use the same active site as dehydrogenase activity (Figure 6B)<sup>43</sup>.

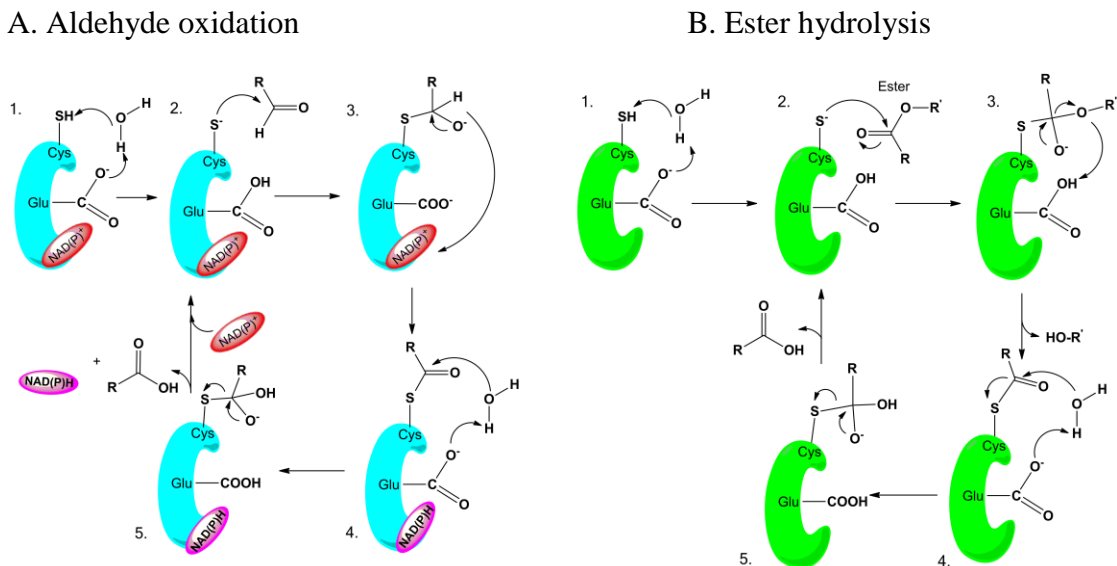


Figure 6: Mechanisms of aldehyde oxidation and ester hydrolysis of ALDH.

A) Aldehyde oxidation occurs in five steps: 1. catalytic cysteine (Cys302) activated by water-mediated proton abstraction by glutamate (Glu268), 2. nucleophilic attack on the carbonyl carbon of the aldehyde by the thiolate group of Cys, 3. formation of a tetrahedral thiohemiacetal intermediate and hydride transfer to the cofactor NAD(P)<sup>+</sup>, 4. Glu activates a water molecule that attacks the carbonyl carbon of the thioester-enzyme complex, 5. formation of tetrahedral intermediate that rearranges to release the carboxylic acid, release of cofactor, and regeneration of activated enzyme by binding cofactor. B) Ester hydrolysis is proposed to use the same active site residues but not require the NAD(P)<sup>+</sup> cofactor. 1. Cys is activated by water-mediated proton abstraction by Glu, 2. thiolate group of Cys attacks the carbonyl carbon of the ester substrate forming an oxyanion intermediate, 3. rearrangement of intermediate results in cleavage and release of an alcohol, 4. Glu<sup>268</sup> activates a water molecule that attacks the carbonyl carbon of the thioester-enzyme complex, 5. formation of tetrahedral intermediate that rearranges to release the carboxylic acid product and activated enzyme with activated cysteine. For step 1 in both reactions, the cysteine may be activated by glutamate directly rather than via an activated water.

Table 2: The 19 Human Aldehyde Dehydrogenase Isoenzymes

Name	% ID to ALDH1A1	Chr Loc	Structure (PDB ID)	Common Substrates / Function	Diseases	Activators (A) / Inhibitors (I)
ALDH1A1	100%	9q21	Sheep (1BXS)	Retinal, Acetaldehyde, DOPAL	Parkinson's disease, cancer, cataracts	DEAB(I), citral(I) disulfiram(I),
ALDH1A2	73%	15q21	Rat (1B19)	Retinal	Cancer, spina bifida	WIN 18446 (I)
ALDH1A3	71%	15q26		Retinal	Cancer	N/A
ALDH1B1	65%	9p11		Acetaldehyde		N/A
ALDH1L1	51%	3q21	ct-Rat (2O2P)	10-formyltetrahydrofolate	Cancer	N/A
ALDH1L2	49%	12q23		Unknown		N/A
ALDH2	68%	12q24	Human (1O02)	Acetaldehyde, 4-HNE, DOPAL	Heart disease, alcoholism	Alda-1(A), daidzin(I), disulfiram(I), citral(I), DEAB(I), benomyI(I)
ALDH3A1	30%	17p11	Human (3SZA)	Aromatic and aliphatic aldehydes	Cancer	Gossypol(I), citral(I)
ALDH3A2	28%	17p11	Human (4QGK)	Fatty aldehydes	Sjögren-Larsson Syndrome	N/A

N/A – no activators or inhibitors available

%ID to ALDH1A1 represents percent protein sequence identity to ALDH1A1 determined by NCBI Blast

Table 2, continued

Name	% ID to ALDH1A1	Chr Loc	Structure (PDB ID)	Substrate / Function	Diseases	Activators (A) / Inhibitors (I)
ALDH3B1	25%	11q13		Unknown	Schizophrenia	N/A
ALDH3B2	27%	11q13		Unknown		N/A
ALDH4A1	28%	1p36	Human (4OE5)	Glutamate $\gamma$ -semialdehyde Proline Catabolism	Cancer; type II hyperprolinemia	N/A
ALDH5A1	34%	6p22	Human (2W8R)	Succinate semialdehyde GABA Metabolism	$\gamma$ -hydroxy-butyric aciduria	N/A
ALDH6A1	30%	14q24	Bacteria (4E4G)	Malonate semialdehyde Valine Catabolism	Metabolic abnormalities	N/A
ALDH7A1	29%	5q31	Human (2J6L)	$\alpha$ -aminoadipic semialdehyde Lysine Catabolism	Cancer	N/A
ALDH8A1	39%	6q23		Retinal		N/A
ALDH9A1	41%	1q23	Cod (1A4S)	$\gamma$ -amino-butyraldehyde Polyamine Metabolism		N/A
ALDH16A1	28%	19q13		Unknown	Gout	N/A
ALDH18A1	29%	10q24	Human (2H5G)	Glutamic $\gamma$ -semialdehyde	Neurologic and metabolic abnormalities	N/A

N/A – no activators or inhibitors available

%ID to ALDH1A1 represents percent protein sequence identity to ALDH1A1 determined by NCBI Blast

The structure of mammalian ALDHs are similar, functioning as multimers, with each subunit containing a catalytic domain, a cofactor binding domain, and an oligomerization domain (Figure 7)<sup>23</sup>. The global structural similarities support identical active site residues whose contributions to aldehyde oxidation are well established. The first crystal structure of an ALDH was determined in 1997 for rat ALDH3A1, a homodimer (PDB Code 1AD3)<sup>44</sup>, with the structure of the first homotetramer solved shortly afterwards for bovine ALDH2 (PDB Code 1AG8)<sup>45</sup>. The first human structure was determined in 1999 (PDB Code ICW3)<sup>46</sup>. To date, the structures of seven human ALDHs have been determined (ALDH2, ALDH3A1, ALDH3A2, ALDH4A1, ALDH5A1, ALDH7A1, and ALDH18A1). The enzyme's active site is in a tunnel that is supported by monomer oligomerization, with the cofactor binding site at one end of the tunnel and the substrate binding site at the other end of the tunnel. Despite similarities in structure and function, the isoenzymes of the ALDH family of proteins have evolved to recognize different spectrums of aldehyde substrates due to differences in the size and shape of their respective substrate binding sites<sup>23,47</sup>.

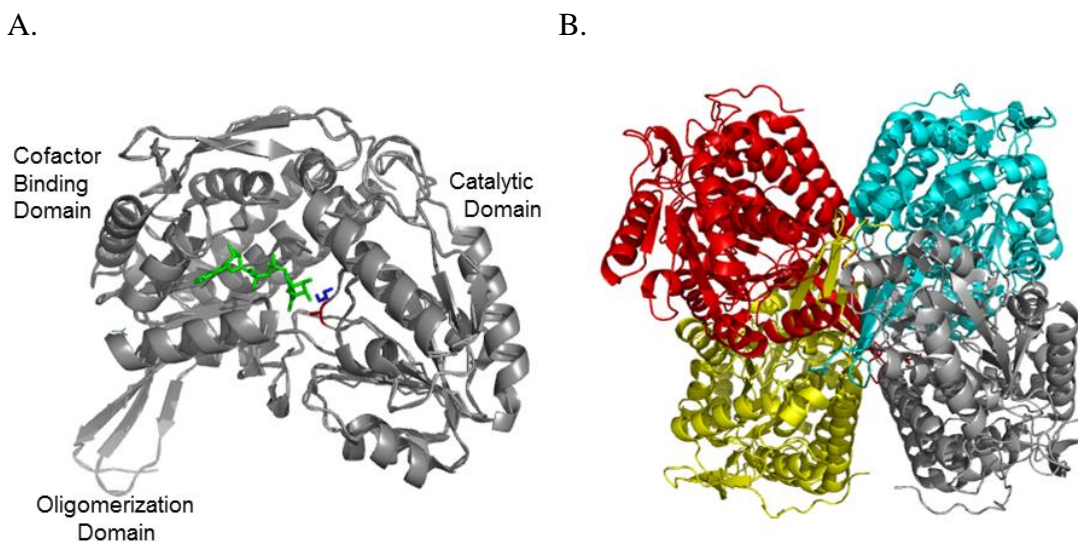


Figure 7: Ribbon diagram of the ALDH2 monomer and homotetramer. A) Monomer showing the cofactor in green, the substrate crotonaldehyde in blue, and the active site Cys-302 in red (PDB Code 1O01). B) ALDH2 homotetramer with each monomer a different color (PDB Code 3N80).



### C. Aldehyde Dehydrogenase 1A1

ALDH1A1 (retinaldehyde dehydrogenase 1, RALDH1) is a highly conserved, cytosolic homotetramer (~55 kDa monomers) that is constitutively expressed in numerous tissues, including brain, liver, kidneys, adipose, eye lens and retina. ALDH1A1 homologs are present in most vertebrate species analyzed, but are absent in Zebra fish and other fishes in the teleost lineage<sup>20</sup>. A key role of ALDH1A1 is the oxidation of retinal (retinaldehyde) to retinoic acid (RA), forming transcriptional regulators critical for normal cell growth and differentiation<sup>48</sup>. Both the substrate (retinal) and product (RA) are important for normal biological processes, including vision, cellular differentiation, and immune function<sup>6</sup>. Other aldehydes oxidized by ALDH1A1 include acetaldehyde during ethanol metabolism, 3,4-dihydroxyphenylacetaldehyde (DOPAL) in dopamine metabolism, and (±)-4-hydroxy-2E-nonenal (HNE), a toxic by-product of oxidative stress. ALDH1A1 also has non-catalytic roles, functioning as structural proteins in eye tissue and as a binding protein for various compounds<sup>22</sup>. In mammalian eye tissue, ALDH1A1, along with other ALDHs, act as corneal and lens crystallins<sup>49</sup>. In this capacity, ALDH1A1 contributes to the transparent and refractive properties of eye tissue, as well as protecting the tissue from ultraviolet radiation damage<sup>22,50</sup>. As a binding protein, ALDH1A1 functions as an androgen-, thyroid hormone-, and cholesterol-binding protein and is also capable of interacting with such drugs as quinolone, daunorubicin, and flavopiridol<sup>22,51,52</sup>. ALDH1A1 shares greater than 70% sequence identity to both ALDH1A2 and ALDH1A3 (RALDH2 and RALDH3, respectively) and both also convert retinal to RA, but their roles may be more confined to embryogenesis and stem cell development<sup>53,54</sup>. ALDH1A1 knockout mice are viable, but ALDH1A2 and ALDH1A3 knockout mice are embryonic or perinatal lethal, respectively, with ALDH1A2<sup>-/-</sup> characterized by heart defects and death at midgestation, and ALDH1A3<sup>-/-</sup> characterized by nasal and ocular defects, leading to respiratory failure shortly after birth<sup>55</sup>. ALDH1A1 can catalyze the oxidation of multiple retinal isomers (all-trans-, 9-cis-, and 13-cis-retinal) with  $K_m$  values ranging from 2 – 6  $\mu\text{M}$ <sup>56</sup>. Both ALDH1A2 and ALDH1A3 prefer the all-trans isomer, but can catalyze the oxidation of the cis-isomers<sup>6</sup>. A fourth retinaldehyde dehydrogenase, ALDH8A1, shares only 40% protein sequence identity with ALDH1A1 and prefers 9-cis-retinal<sup>57</sup>. The resulting retinoic acid products, as well as other retinoids,

act as ligands for the retinoic acid and retinoic X receptors (RAR and RXR, respectively). RAR and RXR are members of the nuclear hormone receptor superfamily, ligand-activated gene regulatory proteins (Figure 8)<sup>58</sup>. The retinoid signaling molecule enters the target cell, binds to the receptor, and activates it. Next, the ligand-receptor complex binds to a specific DNA sequence, the retinoic acid response element (RARE), in the promoter region of a gene, regulating its expression and affecting cellular phenotype. RAR and RXR act as homo- and heterodimers both with other RARs and RXRs as well as other nuclear receptors, including thyroid hormone receptors, peroxisomal proliferator activated receptors (PPAR), and glucocorticoid receptors, enabling great diversity and complexity in the signaling process<sup>59</sup>. Retinoid metabolism and signaling via their nuclear receptors are involved in such fundamental biological processes as embryogenesis, cell cycle control, cell growth and differentiation as well as cellular energy balance. Dysregulation of these functions have been linked to obesity, diabetes, cardiovascular disease, and cancer plus fetal abnormalities and birth defects<sup>29,59</sup>. Although no SNPs in the coding sequence of ALDH1A1 have been associated with a disease phenotype, the National Center for Biotechnology Information (NCBI) AceView

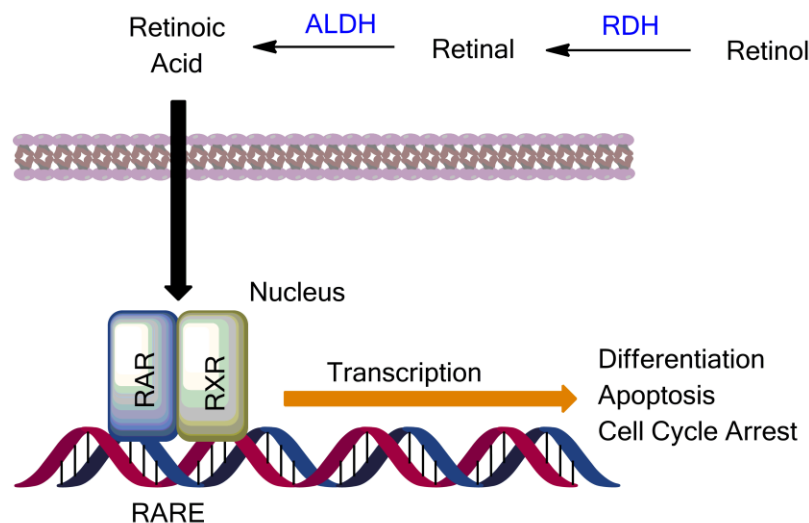


Figure 8: ALDH in the retinoic acid signaling pathway.

Retinol is oxidized by retinol dehydrogenases (RDH) to retinal, which is oxidized to retinoic acid (RA) by the human ALDH isoenzymes ALDH1A1, ALDH1A2, ALDH1A3, and ALDH8A1. RA diffuses into the nucleus, activates transcription factors (RAR, RXR), and initiates transcription of RARE-containing genes. Depending on cellular environment, gene products can lead to differentiation, apoptosis, cell cycle arrest, or other cellular outcomes (modified from Marcato *et al*<sup>58</sup>).

states that there are at least 13 splice variants of ALDH1A1, 5 probable alternative promoter sequences, along with various alternative post-translational modifications sites for polyadenylation, acetylation, methylation, phosphorylation, and ubiquitination<sup>60</sup>. Polymorphisms in the promoter regions have been linked to alcoholism<sup>61,62</sup>. Other diseases have been associated with ALDH1A1 dysregulation, resulting in either too little or too much enzyme, but not to any specific polymorphisms in the gene itself. This dysregulation of ALDH1A1 can result in the disruption of various cellular differentiation pathways due to the enzyme's role in retinoid metabolism. Competition between the oxidation of distinct substrates for ALDH1A1 can also contribute to disease states. As an aldehyde detoxification enzyme, ALDH1A1 is critical for minimizing cellular damage from toxic aldehydes, and a reduction in enzyme activity levels could increase the aldehyde burden in the cell. In addition, ALDH1A1 is involved in the metabolism of different cancer chemotherapeutic agents, and overexpression is correlated with increased drug resistance and poorer prognosis<sup>3,63</sup>. ALDH1A1 is involved in the metabolism of various neurotransmitters, including dopamine, with lowered ALDH1A1 expression potentially leading to the build-up of neurotoxic aldehyde intermediates, including 3,4-dihydroxyphenylacetaldehyde (DOPAL). As will be discussed below, three disorders that have been linked to ALDH1A1 dysregulation are Parkinson's disease, cancer, and obesity.

#### **D. ALDH1A1 and Parkinson's disease**

Parkinson's disease (PD) is a neurodegenerative disorder that affects millions worldwide. Hallmark symptoms of PD include tremors, muscle rigidity, movement and gait abnormalities, with pathological features including the intracellular aggregation of  $\alpha$ -synuclein (Lewy bodies) and the loss of dopaminergic neurons in the substantia nigra<sup>64</sup>. Clinical signs of PD typically do not appear until over 50% of neurons and 80% of striatal dopamine have been lost<sup>65,66</sup>. The neurotransmitter dopamine is primarily found in the central nervous system within two midbrain regions, the substantia nigra and the ventral tegmental area, as well as in the hypothalamus<sup>67</sup>. Dopamine is synthesized in two steps from the essential amino acid tyrosine. Following re-uptake into a neuron, the first step in dopamine catabolism is the production of the aldehyde 3,4-

dihydroxyphenylacetaldehyde (DOPAL) via the enzyme monoamine oxidase (Figure 9). Most DOPAL is oxidized via an ALDH to the carboxylic acid 3,4-dihydroxyphenylacetic acid (DOPAC), but it can also be reduced to the alcohol 3,4-dihydroxyphenylethanol (DOPET). Since the 1960s, PD has been treated primarily with the dopamine precursor levodopa (L-DOPA), managing symptoms by increasing the levels of dopamine in the brain but not halting the progression of the disease. Short-term side effects following levodopa treatment, including nausea and vomiting, are linked to dopamine's role in the peripheral nervous system, while long-term side effects, such as involuntary movements, are likely the result of fluctuations in the body's response to the drug after years of treatment<sup>68,69</sup>. Down-regulation of both ALDH1A1 mRNA and protein has been reported in PD<sup>70-72</sup>, and although DOPAL is a substrate for ALDH1A1, it is also a covalent inhibitor of the enzyme<sup>67,73</sup>. DOPAL has been shown to be cytotoxic in multiple *in vitro* and *in vivo* experiments, with toxicity likely due to protein adduction, isoquinoline formation, and free-radical generation<sup>67</sup>. Low levels of enzyme combined with substrate-induced inhibition of the remaining enzyme would lead to the accumulation of the neurotoxic DOPAL. Both ALDH1A1 and ALDH2 are involved in dopamine metabolism in the brain and double knockout mice (ALDH1A1<sup>-/-</sup> x ALDH2<sup>-/-</sup>) have elevated levels of the neurotoxic aldehydes DOPAL and 4-HNE, have loss of dopaminergic neurons in the substantia nigra, and have age-dependent deficits in motor performance<sup>74</sup>. Therefore, a deficit in ALDH activity, specifically ALDH1A1 activity in the substantia nigra, may lead to the accumulation of neurotoxic aldehydes and ultimately the cell death seen in PD and possibly other neurodegenerative disorders. Activators of ALDH1A1 activity could

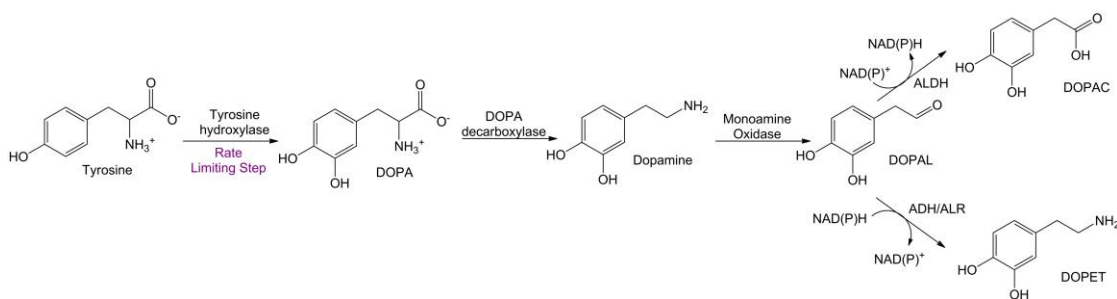


Figure 9: Dopamine synthesis and inactivation.  
DOPA = 3,4-dihydroxyphenylalanine; DOPAL = 3,4-dihydroxyphenylacetaldehyde;  
DOPAC = 3,4-dihydroxyphenylacetic acid; DOPET = 3,4-dihydroxyphenylethanol.

be useful in halting disease progression by either reducing the levels of DOPAL and other toxic aldehydes or by protecting ALDH1A1 from covalent modifications and inhibition by these aldehydes, as seen with the ALDH2 activator Alda-1<sup>75</sup>.

### **E. ALDH1A1 and Cancer**

Over 20 years ago, cytosolic ALDH (ALDH1) was shown to be elevated in hematopoietic stem cells (HSC) and could be used for HSC isolation<sup>76,77</sup>. The first cancer stem cells (CSC) were identified in 1994 for acute myeloid leukemia<sup>78</sup> and nearly a decade later, CSCs in solid tumors were discovered<sup>79</sup>. Since then, CSC have been identified in a number of cancers, and these sub-populations of tumor cells have critical roles in self-renewal, tumorigenesis, and differentiation<sup>80</sup>. Prior to the discovery of CSC, cancer was thought to consist of a heterogeneous collection of cells, each capable of mutating and initiating a new tumor. With the discovery of CSC, the classic stochastic model of cancer was partially replaced with a more hierarchical model. In the hierarchical model for cancer, a small subset of cells are capable of extensive proliferation, while the vast majority of cells in a cancer population are differentiated and have a more limited capacity to spread and form tumors<sup>78</sup>. Unfortunately most therapies still take a stochastic approach to curing cancer, targeting all cells equally and potentially shrinking tumor size in the short-term but allowing CSC to survive and regenerate the cancer<sup>81</sup>. Therefore targeting CSC may offer a better outcome in the quest to eradicate tumors. Expression profiles to determine what genes are up- or down-regulated in stem cells offer potential drug targets for CSC, with various ALDHs having altered expression levels. The roles of ALDH1A1 and other ALDH isoenzymes in stem cells are not well understood. In mice, *Aldh1a1* is expressed in HSC and in neural crest stem cells, but the enzyme is not required for function in either hematopoietic or nervous system stem cells; mice deficient in *Aldh1a1* developed normally, were fertile, and had normal lifespans<sup>82</sup>. However, the mouse genome contains an additional ALDH1A gene, *Aldh1a7*, compared to humans that could compensate for the loss of *Aldh1a1* function<sup>20</sup>. The role of the ALDH1A subfamily members is at least partially due to their involvement in retinoic acid synthesis, as various retinoids, including RA, effect cellular differentiation. All-trans retinoic acid (ATRA) in combination with vitamin D3 induces differentiation followed by cell cycle

arrest in various leukemia cells and as such has been used as an anti-leukemia agent<sup>83</sup>. Inhibition of ALDH via diethylaminobenzaldehyde (DEAB) expands HSC populations by delaying the G<sub>0</sub>G<sub>1</sub> transition<sup>84</sup>, but this effect is reversed upon addition of ATRA, indicating that the role of ALDHs in RA production is important for HSC differentiation<sup>85</sup>. RA has been shown to induce apoptosis and differentiation in a number of human cancer cell models, including oral cancer<sup>86</sup>, non-small cell lung cancer<sup>87</sup>, and liver cancer<sup>88</sup>. The Wnt signaling pathway was first identified as contributing to cancers over a decade before the discovery of CSC, and this pathway is characterized by activation of the transcription factor  $\beta$ -catenin<sup>89</sup>. Mutations in the  $\beta$ -catenin gene itself (*CTNNB1*) typically diminish  $\beta$ -catenin degradation and have been linked to numerous cancers including colorectal, liver, kidney, ovarian, prostate, and thyroid cancers<sup>89</sup>. Targeting and disrupting the Wnt signaling pathway may serve as a potential treatment avenue. Recently, niclosamide, an FDA-approved treatment for tapeworm infections, has been shown to have anti-cancer properties. As an anti-helminthic, niclosamide is proposed to function by uncoupling oxidative phosphorylation in the tapeworm<sup>90</sup>. As an anti-cancer agent, niclosamide inhibits the Wnt pathway and decreases  $\beta$ -catenin levels, along with other signaling cascades<sup>91</sup>. Although initially identified as a candidate for glioblastoma therapy, ovarian cancer cells are particularly sensitive to niclosamide treatment<sup>92</sup>.

First demonstrated in a breast cancer model, *in vitro* assays of mammary stem cells have shown that they form spheroids, or mammospheres, which are rich in progenitor cells and have metastatic capacity<sup>93,94</sup>. These spheroid cells are rich in ALDH1A1 enzyme activity and have been associated with earlier metastasis and poorer clinical outcome in breast cancer<sup>95</sup>. Similar spheroid formation of stem cells has been seen in other cancers, including cervical, prostate, and ovarian cancers<sup>96-98</sup>. Accounting for over half of deaths, ovarian cancer is the most lethal of the gynecological cancers, resulting in nearly 15,000 deaths in the United States in 2011<sup>99,100</sup>. With most patients initially diagnosed in an advanced stage and a reoccurrence rate of 80% following treatment, the 5-year survival rate is only 30%<sup>94</sup>. The ability of ovarian cancer (OC) to form spheroids may facilitate metastases and enable the cancer to survive chemotherapy<sup>101</sup>. Similar to mammospheres,

OC spheroids also have elevated ALDH1A1 enzyme activity, with OC spheroids having up-regulation of both ALDH1A1 and  $\beta$ -catenin transcription, but not elevated transcription of similar ALDHs linked to cancer, including ALDH1A2, ALDH1A3, ALDH2, and ALDH3A1<sup>101</sup>. The ALDH1A1 gene is a direct target of the transcription factor  $\beta$ -catenin, and continuous propagation of OC spheroids results in increasing levels of both ALDH1A1 expression and overall up-regulation of the Wnt signaling pathway from generation to generation of spheroids<sup>101</sup>. The role of ALDH1A1 in spheroid formation is not known, but ALDH1A1 could represent a potential target for drug development to minimize cancer metastasis.

As reviewed in Marcato *et al*, cancer stem cells are often resistant to various chemotherapeutic agents, including daunorubicin, mitoxantrone, temozolomide, carboplatin, paclitaxel, and cyclophosphamide, enabling cancer recurrence<sup>3,58,63,102,103</sup>. As with normal stem cells, CSC are typically in a quiescent state and are not good targets for drugs aimed at fast-growing cells<sup>104</sup>. Increased transporter activity designed to remove drugs from cells and increased DNA repair proteins are also both seen in CSC, decreasing their sensitivity to a drug<sup>58</sup>. In most cases, the role of ALDH1A1 in cancer is not well understood. However, it is known that this enzyme, along with ALDH3A1, metabolizes certain chemotherapeutic agents, including cyclophosphamide, to a less toxic metabolite that can be cleared by the cancer cell, resulting in decreased sensitivity and the need for higher drug dosages<sup>3</sup>. Increases in ALDH expression, particularly ALDH1A1 and ALDH3A1, have been linked to drug resistance in various leukemia cell lines due to inactivation via these enzymes<sup>105</sup>. After progressive rounds of chemotherapy, targeting CSC becomes more difficult as resistance is enhanced. Small molecules that target and inhibit ALDH1A1 activity would be useful as adjuvant therapy for minimizing cancer chemo-resistance. ALDH3A1 inhibitors have been shown to increase sensitivity to mafosphamide, a cyclophosphamide analog that is metabolized by ALDH1A1 and/or ALDH3A1, in cancer cell lines over-expressing ALDH3A1<sup>106,107</sup>. Far less clear is why ALDH1A1 is overexpressed in stem cells in general, and what roles ALDH1A1 plays in maintaining stem cells, self-protection, enhancing metastases, and other phenotypes associated with stem cells and various cancers. Small molecule, chemical probes that

either activate or inhibit ALDH1A1 would be helpful in deciphering the function of the protein and may be developed further as possible therapeutics. As seen with niclosamide, targeting and disrupting the Wnt signaling pathway, including ALDH1A1 expression, may serve as a potential treatment avenue<sup>91,108</sup>.

## **F. ALDH1A1 and Obesity**

A second differentiation pathway regulated by retinoids is adipogenesis. Adipogenesis is the formation of adipocytes, or fat cells, from precursor stem cells. There are two main types of adipose tissue: white adipose tissue (WAT) and brown adipose tissue (BAT). The main function of WAT is the storage of energy as triglycerides, but WAT is also involved in endocrine function, producing adipocyte-derived secreted proteins, adipokines, involved in normal and disease states, including cardiovascular and metabolic disorders<sup>109</sup>. BAT is heavily innervated, rich in mitochondria and uses energy to produce heat via thermogenesis by uncoupling the mitochondrial proton gradient from ATP synthesis<sup>110</sup>. Although BAT is known to be critical for maintaining body temperature in newborns and small mammals, until recently it was not thought to be present in adult humans; its role in adults is not well understood but may be important in energy homeostasis<sup>111-113</sup>. Both types of adipose tissues arise from mesenchymal stem cells, which are also able to differentiate into myocytes, chondrocytes, and osteocytes. As reviewed in Cristancho *et al*, differentiation occurs in two phases, commitment and terminal differentiation<sup>114</sup>. In the commitment phase, pluripotent stem cells are converted to pre-adipocytes, cells indistinguishable morphologically from their precursor yet unable to differentiate into anything other than adipocytes; commitment to BAT or WAT lineages occurs during this first phase of differentiation. During the terminal differentiation phase, pre-adipocytes become mature, functioning adipocytes<sup>114</sup>. The role of adipogenesis in obesity is not well understood; animals fed high fat diets for extended periods first exhibit an increase in adipocyte cell size followed by an increase in adipocyte cell number<sup>115,116</sup>, while rodent models of obesity have increased cell size, increased cell number, or both<sup>117</sup>. In rodents, ALDH1A1 is expressed in adipose tissue, and mice lacking this gene are resistant to both diet-induced obesity and insulin resistance<sup>118</sup>. The ALDH1A1 substrate itself, retinaldehyde, decreased transcription of



adipogenesis genes in cell models of adipogenesis; *in vivo*, retinaldehyde decreased fat levels and increased insulin sensitivity in an obese mouse model, both indicating that retinaldehyde may act as a signaling metabolite inhibiting adipogenesis<sup>118</sup>. ALDH1A1<sup>-/-</sup> mice also had increased body temperature and metabolic rate, plus increased expression of uncoupling protein 1 (Ucp1) and other BAT markers; these results indicate a change in energy expenditure and suggest a possible shift of WAT to more BAT-like characteristics<sup>118,119</sup>. BAT is metabolically active, uses energy to generate heat, and is associated with lower body weight, and a shift from WAT to BAT may reduce obesity<sup>110,119</sup>. The ALDH1A1 product, retinoic acid, also has an effect on adipogenesis. RA treatment of obese mice caused both weight loss and increased insulin sensitivity while increasing expression of RAR and other genes<sup>120</sup>. RAR targets a wide variety of both pro- and anti-apoptotic and cell cycle genes depending on the relative levels of other proteins in the cell. *In vitro* studies have shown that, compared to other vitamin-A metabolizing enzymes used to produce RA, the major enzyme expressed during adipogenesis is ALDH1A1, and adipocytes deficient in ALDH1A1 have impaired adipogenesis<sup>121</sup>. Retinaldehyde, the substrate of ALDH1A1, is an inhibitor of PPAR $\gamma$ , a transcription factor considered the master regulator of adipogenesis<sup>118</sup>. In a negative feedback loop, RA itself may inhibit ALDH1A1 expression<sup>121</sup>. These results indicate that ALDH1A1 may serve as a target for the development of small molecule modulators that could serve both as tools to better understand adipogenesis and as potential therapeutics in the treatment of obesity and related metabolic disorders.

### **G. The Aldefluor Assay**

The controlled regulation of stem cells is critical for normal cell growth and differentiation. In the last few decades, the development of the cancer stem cell hypothesis has increased the urgency to acquire tools to aid in the study of stem cells. As with HSC, ALDH's have been identified as possible biomarkers for numerous CSC<sup>122,123</sup>. A commonly used method to identify and isolate stem cells is via the Aldefluor Assay (Stemcell Technologies, Inc.) that measures ALDH activity in live cells (Figure 10). This assay takes advantage of the conversion of the ALDH substrate BODIPY-aminoacetaldehyde (BAAA) to the charged product BODIPY-aminoacetate, which

accumulates in cells and enhances their fluorescence. The level of fluorescence corresponds to the amount of ALDH activity present in the cell. An inhibitor of ALDH activity, N,N-diethylamino-benzaldehyde (DEAB), is supplied as a negative control for this assay. DEAB was developed as a reversible, competitive inhibitor of ALDH enzymes to replace more toxic, irreversible inhibitors like disulfiram. At the time of its development in the 1980s, DEAB was found to be a potent inhibitor of cytosolic ALDH

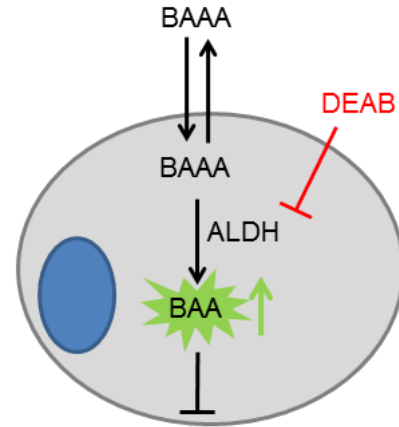


Figure 10: The Aldefluor Assay. Flow cytometry assay that identifies stem cells based on ALDH activity. In the cell, ALDH converts BAAA to BAA, trapping the fluorescent molecule in the cell.

(ALDH1) but not mitochondrial ALDH (ALDH2)<sup>124</sup>. Consequently, the Aldefluor Assay was thought to identify cellular ALDH1A1 activity, suggesting that this isoenzyme was responsible for the vast majority of ALDH activity seen in stem cells. However, recent studies have reported that DEAB inhibits other ALDH isoenzymes; therefore, the Aldefluor Assay will detect stem cells with high levels of other ALDH isoenzyme activity, including ALDH1A2, ALDH1A3, and ALDH2<sup>125-127</sup>. Therefore, ALDH-positive cells identified via the Aldefluor Assay could contain high levels of one or multiple ALDH isoenzymes, and additional assays are needed to determine which isoenzymes are specifically responsible for the elevated ALDH activity. ALDH activity is still an excellent biomarker for multiple normal and cancer stem cell lines, and the Aldefluor Assay is a simple and reliable test for the identification and isolation of these stem cells. However, better control inhibitors, particularly selective inhibitors for each of the ALDH isoenzymes, would aid in understanding which isoenzyme(s) is contributing to the high ALDH activity seen in a particular stem cell. At this time, there are no selective inhibitors for the majority of ALDH isoenzymes, including ALDH1A1.

## H. ALDH1A1 Activators and Inhibitors

The initial focus on the development of ALDH inhibitors was centered on ALDH2 and the development of anti-alcohol (antidipsotropic) therapies due to the enzyme's role in the detoxification of acetaldehyde following alcohol consumption. A single polymorphism in ALDH2 (ALDH2\*2) results in reduced acetaldehyde metabolism and causes alcohol flushing syndrome. Found in close to 50% of the East Asian population, the ALDH2\*2 mutation induces unpleasant side effects, including nausea and erythema, after drinking alcohol, and therefore these populations exhibit lower incidences of alcoholism<sup>128</sup>. By targeting and inhibiting ALDH2, antidipsotropic drugs like daidzin and disulfiram (Antabuse) cause unpleasant side effects following alcohol consumption and discourage users from seeking alcohol. Daidzin is a naturally occurring isoflavone derived from the kudzu plant (*Pueraria lobata*). Although it inhibits both ALDH1A1 and ALDH2, based on IC<sub>50</sub> values, it is 100 times more potent for ALDH2, and daidzin analogs are being developed to better target ALDH2<sup>23</sup>. On the other hand, disulfiram is a potent inhibitor of both ALDH1A1 and ALDH2 and has been used as an antidipsotropic for alcoholism and as a treatment for cocaine addiction, likely due to the role of ALDHs in dopamine metabolism<sup>31</sup>. Disulfiram is also being evaluated for anti-cancer activity resulting from increasing sensitivity to other anti-cancer drugs inactivated by ALDHs as well as by inducing apoptosis<sup>23,129-131</sup>. Citral is a naturally occurring  $\alpha,\beta$ -unsaturated aldehyde found in herbs and citrus fruits that is capable of inhibiting a number of ALDH isoenzymes and likely inhibits by acting as a slow substrate, therefore appearing as an inhibitor for faster substrates<sup>23</sup>. Rodents administered citral have reduced weight gain, lower fasting glucose, and a higher metabolic rate compared to controls, supporting the ALDH1A1<sup>-/-</sup> knockout mice observation that ALDH1A1 may contribute to obesity and other metabolic disorders<sup>132,133</sup>. DEAB, the control inhibitor used in the Aldefluor Assay, is also an aldehyde and likely inhibits by binding at the active site as a substrate that is turned over slowly<sup>127</sup>. Although the effects of disulfiram, citral, and DEAB show promise in targeting ALDH, their ability to inhibit multiple ALDH1/2 family members minimizes their effectiveness as chemical tools, while increasing the likelihood of off-target effects when used as chemotherapeutics. Certain biocides, including the antifungal benomyl and various thiocarbamate herbicides, such as molinate, pebulate, and triallate, also inhibit

various ALDHs, either directly or via biotransformation following exposure<sup>23,134</sup>. Gossypol is a naturally occurring aldehyde that is more selective for ALDH3A1 compared to members of the ALDH1/2 family<sup>23</sup>. Although traditionally used in Chinese medicine as a male contraceptive, its anti-tumor properties are now being studied. Gossypol is not only non-selective compared to other ALDH isoenzymes, it also inhibits other NAD(P)<sup>+</sup>-dependent oxidation enzymes, including aldose reductases and lactate dehydrogenases, and therefore likely interacts with these enzymes at the cofactor binding site<sup>23,135</sup>. Using *in vitro* high throughput screens (HTS), more selective ALDH3A1 inhibitors have been recently developed to sensitize ALDH3A1-positive cancer cells to oxazaphosphorines<sup>106,107,136</sup>. Similar to ALDH3A1, ALDH1A1 is upregulated in multiple cancers, contributing to drug resistance, and therefore inhibitors of ALDH1A1 could increase cancer cell sensitivity to various cancer chemotherapeutics. In a similar manner to ALDH3A1, HTS have been used to find ALDH2 inhibitors, but achieving selectivity for ALDH2 compared to other ALDH enzymes has been more difficult<sup>136,137</sup>. In addition to drug metabolism, ALDH1A1 likely contributes to cancer stem cell maintenance and/or differentiation, and small molecule modulators would aid in better understanding of this enzyme in disease progression as well as serving as possible chemotherapies. The use of RNA interference to silence ALDH1A activity leads to cell cycle arrest, apoptosis, decreased cell viability, and decreased tumorigenesis in human melanoma stem cells<sup>138</sup>. These results indicate that the ALDH1A subfamily represents an excellent therapeutic target for the development of small molecule inhibitors.

There are very few small molecule activators of ALDHs, and the discovery of small molecule enzyme activators, in general, lags far behind inhibitor development<sup>139</sup>. Alda-1 activates ALDH2 and also prevents enzyme inactivation by various toxic aldehyde substrates, including 4-HNE, offering additional protection from oxidative stress<sup>75</sup>. Alda-1 also activates the ALDH2\*2 enzyme, acting as a chemical chaperone to restore function to an otherwise inactive mutant<sup>75,140</sup>. Alda-1 may also activate ALDH1A1 activity following DOPAL-induced inactivation of the enzyme<sup>141</sup>. Similar to inhibitors, enzyme activators can serve as useful tools to decipher the role of an enzyme in various physiological processes. Activators also can have therapeutic value. Alda-1 has been

shown to reduce ischemic damage to the heart following myocardial infarction<sup>75</sup>. Human exposure to pesticides has been linked to Parkinson's disease, and the fungicide benomyl has been shown to inhibit ALDH activity, resulting in the build-up of neurotoxic aldehydes, the death of dopaminergic neurons, and the development of Parkinson's disease<sup>142</sup>. Since down-regulation of both ALDH1A1 mRNA and protein has been reported in PD and a substantial decrease in dopaminergic neurons is needed before manifestation of symptoms, activators of ALDH1A1 could halt or significantly delay the development of PD<sup>65,66,70-72</sup>.

The development of compounds that selectively target ALDH1A1 has proven to be difficult as the ALDH superfamily of enzymes shares many common structural and mechanistic features. These members generally function as homodimers or homotetramers, with the monomers containing three structural domains: a catalytic domain, a cofactor binding domain, and an oligomerization domain. The NAD(P)<sup>+</sup> binding domain is a Rossmann-fold, a nucleotide binding site that consists of two sets of parallel beta sheets and alpha helices. The Rossmann-fold is found in the NAD(P)<sup>+</sup> binding domains of multiple dehydrogenase families, including ALDHs, lactate dehydrogenases, alcohol dehydrogenases, and glyceraldehyde-3-phosphate dehydrogenase<sup>140,143,144</sup>. There are differences in the Rossmann fold between ALDH and other oxidoreductases that could be exploited for the development of small molecule modulators of various ALDH isoenzymes compared to other NAD(P)<sup>+</sup>-binding enzyme families<sup>44</sup>. However, there exists much more structural similarity in the NAD(P)<sup>+</sup>-binding site for ALDH family members, and the development of selective modulators that target this site may be much more difficult. Despite similarities in structure and function, as well as some overlap in substrate preferences, the isoenzymes of the ALDH family of proteins have evolved different aldehyde binding sites due to substitutions of the residues lining their respective substrate binding tunnels<sup>23,47</sup>. Therefore, the substrate binding site or an allosteric site away from the active site would serve as better sites for targeting. The human ALDH1/2 family is a particularly difficult challenge for inhibitor development since it contains the highest number of family members at seven (ALDH1A1, ALDH1A2, ALDH1A3, ALDH1B1, ALDH1L1, ALDH1L2, and ALDH2).

Compounds such as diethylaminobenzaldehyde (DEAB) and disulfiram are potent inhibitors of ALDH1A1, with  $IC_{50}$ 's in the nM range, but both also inhibit ALDH2<sup>23,126</sup>. DEAB is also a relatively potent inhibitor for other ALDH1/2 family members, but is a robust substrate for ALDH3A1<sup>126,127</sup>. DEAB is of particular interest due to its use as a control in the Aldefluor Assay, and further analysis is warranted to better understand the various effects of DEAB on the different ALDH isoenzymes, from potent inhibitor to excellent substrate. In addition, selective ALDH1A1 modulators are needed to better understand this enzyme's role in both normal and disease states. An *in vitro* high throughput screen (HTS) is a common method of discovering novel, small molecule modulators for a particular enzyme. Typically, the rate of aldehyde oxidation by ALDHs is studied by monitoring the formation of NADH at 340 nm on a spectrophotometer (molar extinction coefficient of  $6220 \text{ M}^{-1} \text{ cm}^{-1}$ ) (Figure 11A). However, this approach is not ideal for the screening assay as it is common for compounds in the libraries to absorb light in the same wavelength range as NADH and lead to interference in the analytical approach. Therefore, another assay design is needed for an ALDH1A1 HTS. One

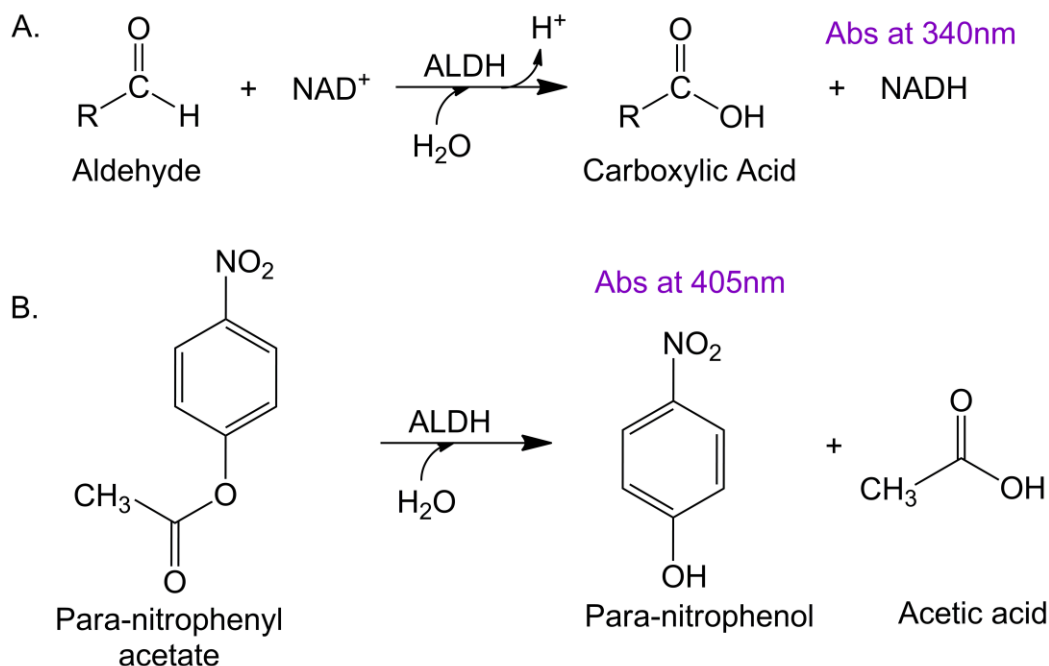


Figure 11: Reactions used to discover ALDH1A1 modulators.

A)  $\text{NAD}^+$ -dependent aldehyde oxidation reaction monitored formation of NADH at 340 nm. B) HTS used an  $\text{NAD}^+$ -independent esterase reaction that monitored the formation of p-nitrophenol at 405 nm.

possibility is to couple aldehyde oxidation to a second reaction that can be monitored by either fluorescence or UV/Vis spectrophotometry. For example, the dehydrogenase activity of ALDH2 was coupled to the NADH-dependent reduction of resazurin to resorufin to discover Alda-1, a novel activator of ALDH2<sup>75</sup>. As ALDH1A1 can also function as an esterase, another approach would be to use the inherent esterase activity of ALDH1A1 to identify modulators. The ALDH1A1 ester substrate para-nitrophenylacetate (pNPA) is hydrolyzed to p-nitrophenol, which absorbs light at 405 nm and can be monitored spectrophotometrically, with minimal interference from library compounds (Figure 11B). Site-directed mutagenesis studies on ALDH2 have shown that Cys-302 is the essential nucleophile for both the dehydrogenase and the esterase reaction, with Glu-268 acting as the general base to activate Cys-302<sup>38,39</sup>. The proposed catalytic steps for both the dehydrogenase and esterase reactions have been recently reviewed<sup>23</sup> and are shown in Figure 6, although minor details still need to be resolved, including the roles of second sphere residues in assisting proton transfer to solvent<sup>145,146</sup>. The use of common active site residues for the two reactions makes it likely that modulators of the esterase reaction would also modulate aldehyde oxidation activity. In support of this hypothesis, the ALDH2 activator Alda-1 activates both the esterase and dehydrogenase activity of the enzyme, and daidzin inhibits both reactions<sup>75,140,147</sup>. An additional advantage of the esterase reaction is that it does not require the cofactor NAD<sup>+</sup>, and so allows the screen to be less influenced by compounds binding to this site. Small molecules that alter ALDH1A1's esterase activity can be further analyzed using steady-state kinetics, protein X-ray crystallography, and cell culture studies to determine the compound's effect on the aldehyde oxidation activity of ALDH1A1 and other ALDH superfamily members, their mechanism of action, and ultimately determine their usefulness as both chemical tools and chemotherapeutics.

## **I. Hypothesis and Approach**

Due to differences in the structure and function of human ALDH1A1 isoenzyme, we hypothesize that small molecules can be identified that selectively modulate the activity of ALDH1A1 compared to other members of the ALDH superfamily. Comparisons of the active sites of sheep ALDH1A1 and human ALDH2 and ALDH3A1 illustrate that there

are sufficient differences to support the development of inhibitors specific for each enzyme (Figure 12). The substrate binding site of sheep ALDH1A1 (PDB 1BXS)<sup>148,149</sup> has a large, somewhat linear active site, while human ALDH2 (PDB 1O02)<sup>150</sup> has a more narrow, linear active site. In contrast, human ALDH3A1 (PDB 3SZA)<sup>137</sup> has a large, bent active site.

ALDH1A1 has been linked to both normal and diseased states. ALDH1A1 is one of four retinaldehyde dehydrogenases that produce retinoic acid, a key transcription regulator in a number of cell growth and differentiation pathways. Up-regulation of the gene and increased protein levels are seen in both normal and cancer stem cells. Increased ALDH1A1 protein levels are often correlated with increased drug resistance, more aggressive tumors, and poorer prognosis. Rodent studies using either gene knockout or non-selective inhibition indicate that ALDH1A1 is involved in adipogenesis and obesity. Down-regulation of the enzyme is observed in Parkinson's disease, with build-up of toxic substrates of ALDH1A1 possibly contributing to PD development. Therefore, ALDH1A1 is a potential target for drug therapy. However, the specific role ALDH1A1 is playing in many of these biological processes and disease states is unclear. Is ALDH1A1 functioning via retinaldehyde metabolism, another aldehyde substrate, or is it a completely different function of the enzyme? What contribution does ALDH1A1 play compared to other aldehyde dehydrogenases, especially the other retinaldehyde dehydrogenases? The discovery of small molecules that selectively activate or inhibit

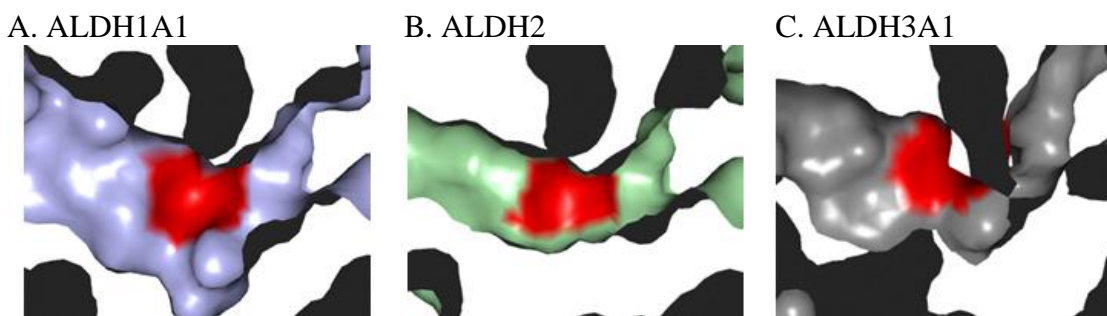


Figure 12: Comparison of the active sites of three ALDH isoenzymes. The surface rendition of the active site, with the cysteine shown in red and the cofactor binding site orientated to the right of the cysteine. A) Sheep ALDH1A1 (PDB 1BXS); B) Human ALDH2 (PDB 1O02); C) Human ALDH3A1 (PDB 3SZA).



ALDH1A1 aldehyde oxidation activity but have little to no effect on other ALDHs could serve as chemical tools to better delineate the biological function(s) this enzyme has in both normal and disease states. With better understanding of the physiological processes involved, it is possible to accurately determine potential targets for drug development, maximizing therapeutic benefits while minimizing side effects (Figure 13)<sup>151</sup>.

With the knowledge that ALDH1A1 is involved in a number of biological pathways, our goal was to characterize small molecule inhibitors of ALDH1A1 that can serve as chemical tools to decipher the specific function of this protein. At this time, there are no commercially available selective inhibitors of ALDH1A1 and therefore no chemical tools capable of distinguishing the role of ALDH1A1 compared to a number of structurally and functionally similar ALDH isoenzymes. Although RNAi can enable specific knockdown of each of the ALDH isoenzymes, small molecule probes do offer some advantages. First, a good modulator (activator or inhibitor) will modify the enzymatic activity of a protein but have little to no effect on other functions of that protein. Second, activation or inhibition by small molecules is dose dependent and often reversible, enabling a high level of control for the pathway of interest. Finally, enzyme activation studies are not possible using RNAi. Ideally, chemical probes and RNAi should complement each other, possibly revealing any off-target effects of the compound and/or nonenzymatic roles of the protein of interest.

We took two approaches to develop selective activators and inhibitors of ALDH1A1:

1. the use of a nonselective inhibitor as a scaffold to develop more selective inhibitors,

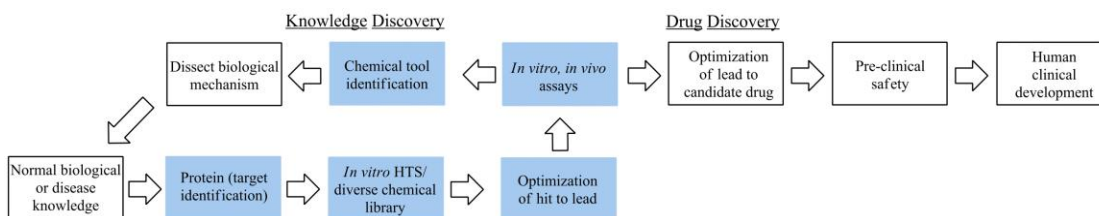


Figure 13: Small molecule discovery.

Approach to the discovery of small molecules as chemical tools for scientific discovery and for drug development. The steps in the blue boxes were performed as part of this project (modified from Lipinski *et al*<sup>151</sup>).

and 2. the discovery of novel compounds from a high throughput screen. To accomplish our goal, first we characterized DEAB, a commonly used, commercially available, non-selective inhibitor of ALDH1A1 and proposed a mechanism of action for its varied effect amongst the ALDH isoenzymes, from a pseudo-covalent inhibitor to a rapid substrate. By determining why it behaves differently within the active sites of different ALDHs, chemical modification to DEAB could produce more selective inhibitors of ALDH isoenzymes. Second, we used X-ray crystallographic studies to determine the structure of human ALDH1A1, enabling comparisons between other human ALDHs and identifying potential sites that can be exploited for compound selectivity. Finally, an *in vitro* esterase high-throughput screen was developed to identify novel activators and inhibitors of ALDH1A1. Compounds identified in the HTS were characterized using steady-state enzyme kinetics and X-ray crystallographic studies. Ideally, the knowledge gained from the discovery of these compounds will enable the development of new medicines to treat disease.

## II. Materials and Methods

### A. Materials

Chemicals and reagents used for protein expression and purification, enzyme kinetics, and X-ray crystallography were purchased from Sigma Aldrich (St. Louis, MO), unless otherwise noted. Compounds were purchased from ChemDiv Corporation (San Diego, CA) and through the Indiana University Chemical Synthesis Core. Dithiothreitol (DTT) was purchased from GoldBio (St. Louis, MO). Primers for site-directed mutagenesis and sequencing were purchased through Integrated DNA Technologies (Coralville, IA). Sitting drop plates for crystallography were purchased from Charles Supper Co (Natick, MA). Sequencing was performed by the Department of Biochemistry and Molecular Biology, Indiana University School of Medicine.

### B. Methods

#### 1. Production of ALDH Isoenzymes

The full-length cDNA for human ALDH1A1 was generously provided by Dr. Henry Weiner, Purdue University in a pT7-7 vector<sup>152</sup> and used to transform *E. coli*. BL21 (DE3) cells. The general protocol for expression and purification of ALDH1A1 has been published before and is similar to ALDH2<sup>136,152</sup>. Specifically, a single transformed BL21 colony was first grown in 10 mL 2XTY broth containing 100 µg/mL ampicillin (TY-Amp) for 8 – 10 hours at 37°C with shaking. Next, 5 mL of this starter culture was used to inoculate 100 mL TY-Amp and grown overnight at 37°C with shaking. For large-scale preparations (8L), 8 x 1L TY-Amp in baffled 2L flasks were inoculated with 5 mL of the overnight culture and incubated at 37°C with shaking at 200 rpm. When the culture reached an optical density at 600 nm (OD<sub>600</sub>) of 0.6 to 0.8, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 500 µM. IPTG is used to induce large quantities of ALDH1A1 protein expression via the T7 promoter on the plasmid<sup>153</sup>. Cultures were incubated overnight at 16°C with shaking at 200 rpm before harvesting. The cells were centrifuged to remove broth, and the pellets flash frozen in liquid nitrogen and stored at -80°C. For ALDH expression and purification, all buffers contained dithiothreitol (DTT), which helps maintain cysteine residues in their reduced state (-SH), and were sparged with helium to remove dissolved oxygen.

Table 3: Solutions used in the preparation of ALDH1A1 protein.

	Lysis Buffer	DEAE Elute Buffer	HAP Elute Buffer
Sodium Phosphate, NaH <sub>2</sub> PO <sub>4</sub>	10 mM, pH 7.0	10 mM, pH 7.0	20 mM, pH 7.5
EDTA	2 mM	2 mM	1 mM
Benzamidine	1 mM	1 mM	---
DTT	1 mM	1 mM	1 mM
NaCl	---	250 mM	50 mM
4-HAP	---	---	10 mM

Combined, sparging and DTT minimized oxidation of active site cysteine residues. The cell pellets were thawed slowly in a cold water slurry and resuspended in lysis buffer (Table 3). Cells were lysed via three passages through a microfluidizer (DivTech Equipment). Cell lysate was clarified via centrifugation at 35,000 rpm for 30 minutes at 4°C (Beckman-Coulter Optima L-90K Ultracentrifuge, Ti-45 Rotor). The clarified lysate was dialyzed into two 4L changes of lysis buffer at 4°C, each at least 5 hours. Next, the lysate was loaded onto a DEAE-Sepharose column and protein eluted with a gradient of 0 to 250 mM NaCl (DEAE Elute Buffer, Table 3) with most elution occurring at approximately 100 mM NaCl. The eluted fractions from the DEAE-Sepharose column were analyzed by SDS gel to confirm the presence of protein (Figure 14). Fractions containing protein were pooled and loaded onto a 4-hydroxyacetophenone (4-HAP) affinity column and eluted in a single step with approximately 75 mL of 10 mM 4-HAP elute buffer (Table 3, Figure 14). The protein was dialyzed exhaustively into three 4L changes of 10 mM sodium ACES, pH 6.6 with 1 mM DTT at 4°C, each at least 5 hours. Protein was concentrated to 5 – 10 mg/mL using Amicon Ultra Centrifugal Devices (Millipore Corp, Bedford, MA). Prior to storage, protein concentration was measured using the BioRad Protein Assay (BioRad Laboratories, Hercules, CA), and its aldehyde oxidation activity determined using saturating amounts of propionaldehyde and NAD<sup>+</sup>. An 8L preparation generated 150 – 200 mg ALDH1A1 protein. The majority of the protein was concentrated to 5 – 8 mg/mL, distributed into 100 – 200 µL aliquots, flash

frozen in liquid nitrogen, and stored at  $-80^{\circ}\text{C}$ . ALDH1A1 for crystallization was stored at  $-20^{\circ}\text{C}$  in a 50% (v/v) solution with glycerol to minimize degradation and lose of activity; prior to use it was dialyzed into three 4L changes of 10 mM sodium ACES, pH 6.6 with 1 mM DTT at  $4^{\circ}\text{C}$ , each for at least 5 hours.

The ALDH1A1 cDNA contained a known A-to-G SNP at position 72928972 on chromosome 9 (NCBI rs1049981) resulting in an Asn-to-Ser missense mutation at protein position 121<sup>152</sup>. This SNP has been found in a small percentage of the HapMap-CEU population representing Utah residents with Northern and Western European ancestry, but there is no known clinical significance to the mutation. ALDH1A1 N121S was active and behaved similar to ALDH1A1 WT. ALDH1A1 WT protein was used for all aldehyde oxidation assays and X-ray crystallography, unless otherwise noted. The full length cDNA for human ALDH4A1 and ALDH5A1 were generously provided by Dr. Daria Mochly-Rosen. ALDH1L1 (10-formyltetrahydrofolate dehydrogenase) consists of three domains: an N-terminus hydrolase domain, an intermediate linker domain, and the C-terminus ALDH domain<sup>154</sup>. The carboxyl terminus of rat ALDH1L1 was generously

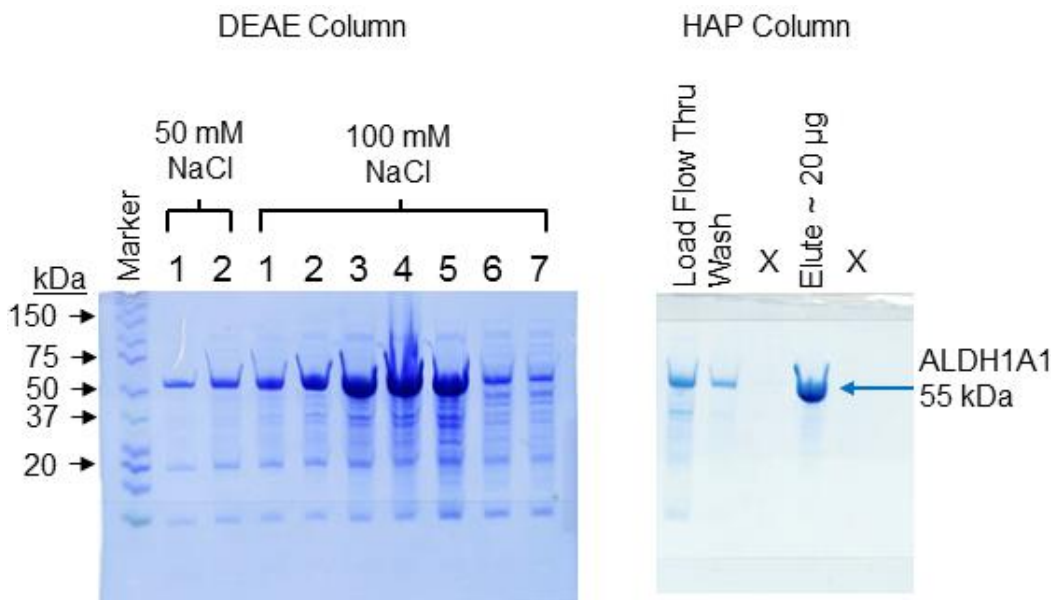


Figure 14: ALDH1A1 protein purification. Representative SDS-PAGE gels of WT ALDH1A1 protein expression and purification. From the DEAE column, 50 mM NaCl fractions 1 and 2, plus 100 mM NaCl fractions 1 – 5 were pooled and loaded on to the HAP column.

provided by Dr. Sergey Krupenko in the pRSET expression plasmid. ALDH1L1 was expressed and purified as previously described for ALDH3A1<sup>136</sup> with the following modifications: 1) the medium contained 100 µg/mL ampicillin, 2) cells were lysed via 3 passages through a microfluidizer (DivTech Equipment), and 3) a single passage on a nickel-NTA column was used for purification, without the second Q-sepharose column used to purify ALDH3A1. Protein used for kinetics was flash frozen in liquid nitrogen and stored at -80°C. ALDH1A2, ALDH1A3, ALDH1B1, ALDH2, and ALDH3A1 were produced and purified as previously described<sup>107,136,155</sup>. ALDH3A1 was purified and provided by Dr. Bibek Parajuli; ALDH1A2, ALDH1A3, ALDH1B1, and the mutant ALDH1A1 G458N were purified and provided by Lanmin Zhai; ALDH4A1 and ALDH5A1 were purified and provided by Cameron Buchman similar to the purification of ALDH1L1.

## **2. Generation of wild-type ALDH1A1 and the G458N mutant**

The QuikChange site-directed mutagenesis protocol was used to make point mutations. First, wtALDH1A1 was generated from the Weiner N121S polymorphism. Wild-type ALDH1A1 was constructed using the forward primer 5'- CTC TAT TCC AAT GCA TAT CTG AAT GAT TTA GCA GGC TGC ATC -3' and its complement. A point mutation of wtALDH1A1 was performed to produce the G458N mutant. ALDH1A1 G458N mutants were constructed using the forward primer 5'-GTG GGT GAA TTG CTA TAA CGT GGT AAG TGC CCAG-3' and its complement. This G458N mutant was purified in the same way as other ALDH1A1 proteins (WT, N121S), but the yield was approximately 10% the yield of WT. G458N was stored at 2-mg/mL and 8-mg/mL at -80°C. Kinetic experiments for G458N were performed in the same manner as WT protein.

## **3. X-ray crystallography**

ALDH1A1 protein used for crystallization was stored at -20°C in a 50%(v/v) glycerol solution. Prior to use, the protein was dialyzed against three 4L changes of 10 mM sodium ACES, pH 6.6 and 1 mM dithiothreitol at 4°C, each for at least 5 hrs. Hampton Research (Laguna Niguel, CA) screens, in particular their Index and Additive Screens,

were used to help determine appropriate crystallization conditions for the human ALDH1A1 protein. Other than the Hampton Research screens and the compounds used for co-crystallization or soaks, all solutions (buffers, NAD<sup>+</sup>/NADH, and protein) were sterile-filtered thru a 0.2  $\mu\text{m}$  or 0.45  $\mu\text{m}$  membrane to remove any particulate matter. Crystallization conditions were determined using a sitting drop method by mixing 2  $\mu\text{L}$  of the protein solution (4 mg/mL) with 2  $\mu\text{L}$  of crystallization solution, with a 500  $\mu\text{L}$  reservoir solution. The following Hampton Research Screens were initially tested to determine what general conditions were most favorable for ALDH1A1 crystallization: Index (HR2-144), Crystal Screen (HR2-110), Crystal Screen 2 (HR2-112), Polyethylene Glycol 6000 (HR2-213), Sodium Chloride (HR2-219), PEG/Ion (HR2-126), and PEG/Ion 2 (HR2-098). Promising results from these screens were further tested using the Additive Screen (HR2-428). For the Additive Screen, 0.2  $\mu\text{L}$  of the additive screen component was added to the sitting drop, using the conditions identified as favorable from previous screens.

Following optimization, the final crystallization solution contained 100 mM sodium BisTris, pH 6.2 – 7.5, 7 – 12% polyethylene glycol 3350 (PEG3350 from Hampton Research, Laguna, CA), 200 mM sodium chloride, and 5 – 10 mM ytterbium chloride. Protein (3 – 7 mg/mL) was crystallized primarily using the sitting drop method, but the hanging drop method also successfully produced crystals under the same conditions. Drop composition could also vary, but an equal ratio of protein volume to reservoir volume was preferred, and typically 2 – 3  $\mu\text{L}$  protein was mixed with an equal volume of crystallization solution. The drop was mixed by both pipetting and agitation. Crystallization was allowed to proceed at 22°C in an incubation chamber. Crystals also formed when placed in a 10°C incubation chamber. Small crystals were present within 24 hours, while larger crystals were present in 4 – 7 days. Crystals continued to diffract well for 6 weeks, but beyond that timeframe, the crystals became soft, malleable, and no longer diffracted well. For the ALDH1A1-NADH structure, seven day old apo-crystals were soaked for 2 hours with crystallization solution containing 1 mM NAD<sup>+</sup>. Complexes with CM026 analogs were produced by soaking apo-crystals overnight with a crystallization solution containing 500  $\mu\text{M}$  inhibitor and 2% DMSO. Complexes with

CM037 analogs were produced by soaking apo-crystals for 5 hours with a crystallization solution containing 500  $\mu\text{M}$  CM037, 1 mM  $\text{NAD}^+$ , and 1% DMSO. Freezing of the crystals occurred in crystallization solution plus 20% (v/v) ethylene glycol as well as any soak components ( $\text{NAD}^+$  or compound). Crystals were frozen in a liquid nitrogen cryostream, and diffraction tested on the home source Bruker D8 System. Crystals with good diffraction were stored in liquid nitrogen until collection. Diffraction data was collected at either Beamline 19-ID or 19-BM operated by the Structural Biology Consortium at the Advanced Photon Source, Argonne National Laboratory, Chicago, Illinois. Diffraction data was indexed, integrated, and scaled using either HKL2000 or HKL3000 program suites<sup>156</sup>. The CCP4 program suite<sup>157</sup> was used for molecular replacement and refinement. For molecular replacement for the apo-ALDH1A1 structure, sheep ALDH1A1 (PDB Code 1BXS) was used; for all other structures, the apo-ALDH1A1 was used. The molecular graphics application Coot<sup>158</sup> was used for model building, and the TLSMD (Translation/Libration/Screw Motion Determination) server was used to determine dynamic properties of the protein<sup>159,160</sup>.

#### 4. Esterase High-Throughput Screen

The high-throughput screen (HTS) was performed using an esterase reaction (Figure 15). The esterase assay was selected for screening because it is cofactor independent and the spectral properties of the monitored product (para-nitrophenol at 405 nm) are less likely to overlap with the absorbance characteristics of the compounds in the libraries. The

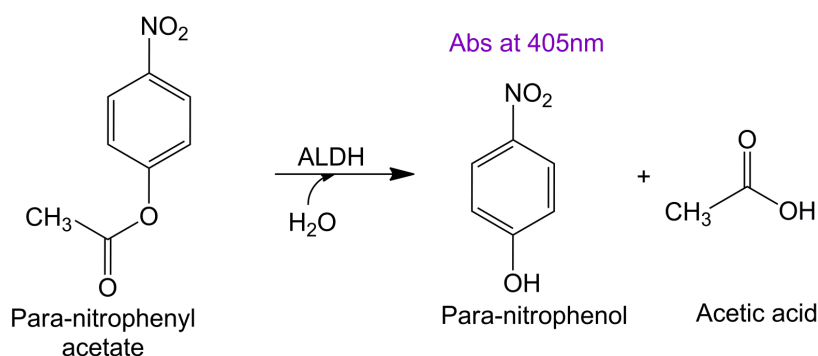


Figure 15: Esterase high-throughput screen.

The high-throughput screen used an  $\text{NAD}^+$ -independent esterase reaction that monitored the formation of para-nitrophenol at a wavelength of 405 nm.



substrate was prepared by adding 0.0725 g of para-nitrophenylacetate (pNPA) to 6 mL of DMSO; this mixture was slowly added to 94 mL of water with continuous stirring for a final stock concentration of 4 mM in 6% DMSO. ALDH1A1 was diluted to a concentration of 3.6  $\mu$ M in 10 mM sodium HEPES, pH 7.5 and maintained on ice until added to the reaction plate. The non-selective ALDH1A1 inhibitor Aldi-1<sup>137</sup> was used as a positive control for inhibition, with a stock concentration of 62.5  $\mu$ M in a 2% DMSO solution and a final reaction concentration of 25  $\mu$ M. To identify novel activators and inhibitors of ALDH1A1, a HTS of 64,000 compounds from ChemDiv Corporation (San Diego, CA) was carried out at the Indiana University Chemical Genomics Core Facility. This library contained a structurally diverse collection of drug-like compounds that follow Lipinski's rule<sup>151</sup>. The HTS was performed in 384-well, clear-bottomed plates, monitoring the change in absorbance of p-nitrophenol at a wavelength of 405 nm. The final 50- $\mu$ L reaction contained 730 nM ALDH1A1, 800  $\mu$ M substrate para-nitrophenylacetate (pNPA), 10  $\mu$ M library compound, and 2% DMSO in 25 mM sodium HEPES, pH 7.5 at 25°C. The rate of reaction was determined by monitoring the reaction for 7 minutes, with prior optimization results indicating that the reaction is linear for at least 20 minutes. Prior to conducting the HTS, a Z factor was calculated by comparing the values of ALDH1A1 plus/minus the control inhibitor Aldi-1 under the conditions of the HTS assay, each at n = 384, using the following formula:

$$Z \text{ factor} = 1 - \frac{3 \left| \delta_p - \delta_n \right|}{\left| \mu_p - \mu_n \right|}$$

where  $\mu_p$  and  $\delta_p$  are the mean and standard deviation for the control (positive) reaction, respectively, and  $\mu_n$  and  $\delta_n$  are the mean and standard deviation for the inhibition (negative) reaction, respectively. A Z factor between 0.5 and 1.0 indicates a good assay that is statistically strong enough to detect inhibitors.

The IU Chemical Genomics Facility provided plates containing 20  $\mu$ L aliquots of each compound at a 25  $\mu$ M concentrations in a 2% DMSO solution, for a final concentration of 10  $\mu$ M in the 50  $\mu$ L screening reaction. Columns 23 and 24 contained no compound and were used as assay controls. A sample reaction plate is shown in Figure 16, with columns 1 – 22 containing library compounds and columns 23 and 24 used as controls:

column 23 as a control for enzyme activity (positive control) and column 24 as an inhibition control (negative control). The enzyme activity control formed the basis to determine whether and to what degree a library compound had an effect on enzyme activity. The control for inhibition combined with the enzyme control established whether that plate was run successfully. If the values in these two columns were not significantly different, an error had occurred in set-up, or enzyme activity had decreased too much to

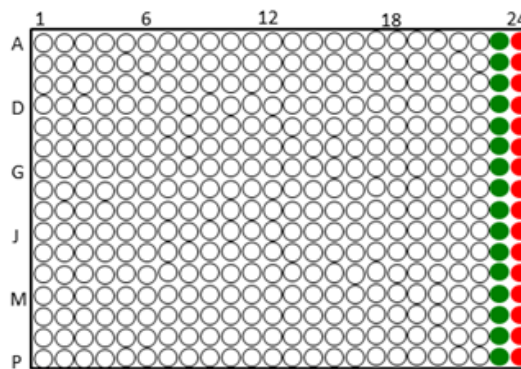


Figure 16: Esterase HTS plate set-up. High-throughput screen 384-well plate. The open circles represent wells containing library compounds. Green circles are the enzyme control and red circles are the inhibition control.

identify inhibitors. For the two control columns, a multichannel pipet was used to add either 20  $\mu\text{L}$  of a 2% DMSO solution or 20  $\mu\text{L}$  of the control inhibitor Aldi-1 in a 2% DMSO solution. To each plate, 10  $\mu\text{L}$  of 125 mM HEPES buffer was added via the Titan Multidrop-384 Microplate Dispenser. The Tecan Genesis or the Tecan Freedom EVO Robotic workstations were used to add 10  $\mu\text{L}$  of a 3.6  $\mu\text{M}$  ALDH1A1 enzyme solution followed by 10  $\mu\text{L}$  of a 4 mM pNPA solution in 6% DMSO with mixing. The plates were centrifuged at 1000 rpm for 10 seconds and read on a Spectromax Plus 384 plate reader over a 7-minute period. The rate of reaction was used to determine a ratio to control for each compound, with control being the average of the rates of reaction for the plate control (positive control), column 23 ( $n = 16$ ):

$$\text{Ratio to Control} = \frac{\text{Rate of reaction with compound}}{\text{Average rate of reaction of control}}$$

A compound was identified as an activator if its ratio to control was greater than 2 and was identified as an inhibitor if its ratio to control was less than 0.4. Compounds identified as either activators or inhibitors of ALDH1A1 esterase activity were rescreened using the same protocol and cutoffs to confirm the initial readings. Confirmed hits were grouped based on their structural properties. Compounds selected for further analysis

were purchased from ChemDiv Corporation (San Diego, CA) to determine their effect on aldehyde oxidation.

## 5. Aldehyde Oxidation Activity Assays for ALDH Isoenzymes

Hits from the HTS were ordered from ChemDiv to determine if they had an effect on aldehyde oxidation and if they were selective for ALDH1A1 compared to ALDH2 and ALDH3A1. The dehydrogenase activity of the three isoenzymes was assayed by monitoring the production of NADH at 340 nm (molar extinction coefficient of  $6220 \text{ M}^{-1}\cdot\text{cm}^{-1}$ ) on a Beckman DU-640 or Cary 300 Bio UV-Vis spectrophotometer for 2 to 3 minutes. Under these conditions, the control reactions were linear for at least 10 minutes. For ALDH1A1 and ALDH2, the reaction contained 100 – 200 nM enzyme, 200  $\mu\text{M}$   $\text{NAD}^+$ , 100  $\mu\text{M}$  propionaldehyde, and 1% DMSO in 50 mM sodium BES, pH 7.5 at room temperature (Figure 17A). For ALDH3A1, the reaction contained 20 – 25 nM enzyme, 200  $\mu\text{M}$   $\text{NAD}^+$ , 300  $\mu\text{M}$  benzaldehyde, and 1% DMSO in either 100 mM sodium phosphate or 50 mM sodium BES at pH 7.5 at room temperature (Figure 17B). For most compounds, a concentration of 20  $\mu\text{M}$  was used for the initial round of selectivity assays. However, due to solubility issues for CM307, 10  $\mu\text{M}$  of compound

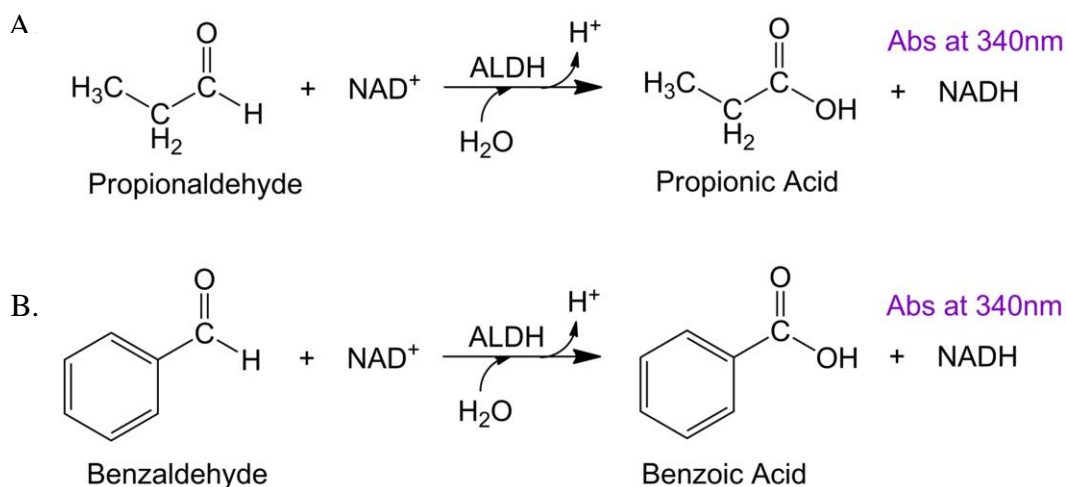


Figure 17: Aldehyde oxidation activity assays.  $\text{NAD}^+$ -dependent aldehyde oxidation reaction used to determine selectivity of hit. Both reactions monitored the formation of NADH at 340 nm. A) oxidation of propionaldehyde used with ALDH1 family members, B) oxidation of benzaldehyde used for ALDH3A1.

was used. Following a 2 minute incubation of enzyme, compound, and  $\text{NAD}^+$ , the reaction was initiated by adding substrate. Compounds that were selective for ALDH1A1 compared to ALDH2 and ALDH3A1 were tested on other ALDH1 isoenzymes. For ALDH1A2, ALDH1A3, and ALDH1B1, the reaction contained 100 – 200 nM enzyme, 200  $\mu\text{M}$   $\text{NAD}^+$ , and 100  $\mu\text{M}$  propionaldehyde. Selectivity towards ALDH1L1 was tested with 500 nM enzyme, 4 mM propionaldehyde, and 500  $\mu\text{M}$   $\text{NADP}^+$ . Selectivity towards ALDH4A1 was tested with 100 nM enzyme, 20 mM propionaldehyde, and 1.5 mM  $\text{NAD}^+$ . Selectivity towards ALDH5A1 was tested with 100 nM enzyme, 2 mM propionaldehyde, and 1.5 mM  $\text{NAD}^+$ . For all selectivity assays, reactions contained 1% DMSO in 50 mM sodium BES, pH 7.5 at room temperature. Following a 2 minute incubation of enzyme, compound, and  $\text{NAD(P)}^+$ , the reaction was initiated by adding propionaldehyde, and the production of  $\text{NAD(P)H}$  was monitored at 340 nm on either a Beckman DU640 or Cary 300 Bio UV-Vis Spectrophotometer for 2 – 3 minutes.

For compounds with over 60% inhibition at 20  $\mu\text{M}$ ,  $\text{IC}_{50}$  values for propionaldehyde oxidation were calculated by varying the concentration of the compound from 0 – 200  $\mu\text{M}$ . Following a 2 minute incubation of enzyme, compound, and  $\text{NAD}^+$ , the reactions were initiated by adding propionaldehyde and monitored at 340 nm for 2 – 3 minutes. Data were fit to the four parameter  $\text{EC}_{50}$  equation using SigmaPlot (StatSys v12.3). The values represent the mean/SEM of three independent experiments, with each experiment  $n = 3$ .

## **6. Steady-State Enzyme Kinetics**

In order to use steady-state enzyme kinetics to characterize these compounds, we first needed to determine the  $K_m$  values of the aldehyde substrates and cofactor  $\text{NAD}^+$  for ALDH1A1. Acetaldehyde was distilled at room temperature to remove impurities, and the concentration of acetaldehyde in the distillate was calculated using ALDH2. Acetaldehyde has a sub-micromolar  $K_m$  for ALDH2 and is rapidly turned over, producing  $\text{NADH}$  at a constant rate until all acetaldehyde has been metabolized. At this time, the velocity curve flattens to a slope of zero, as no more  $\text{NADH}$  is being produced. The amount of acetaldehyde present at the start can be calculated from the total amount of

NADH produced. This method to determine the total acetaldehyde present was used for all reactions requiring varied concentrations of acetaldehyde. The reactions for  $K_m$  calculations contained 100 nM ALDH1A1 in 50 mM sodium BES at pH 7.5 at room temperature and were monitored at 340 nm for 3 minutes. To determine the  $K_m$  for the aldehyde substrates, we produced velocity curves, plotting increasing concentrations of aldehyde (0 to 2 mM) versus the velocity of the reaction, with a saturating amount of  $\text{NAD}^+$  (1.5 mM). The  $K_m$  value is the aldehyde concentration where velocity is one-half its maximum value. A similar method was used to calculate the  $K_m$  for cofactor, plotting increasing amounts of  $\text{NAD}^+$  (0 – 1.5 mM) versus velocity, with a saturating amount of propionaldehyde (2 mM). Data were plotted using the Simple Ligand Binding Toolbox in SigmaPlot (StatSys v12.3), and the values represent the mean/SEM of three experiments (each  $n = 3$ ).

Characterization of ALDH1A1 WT and ALDH1A1 G458N were performed by co-varying acetaldehyde and  $\text{NAD}^+$  concentrations in reactions containing 150 – 300 nM enzyme, 50 – 500  $\mu\text{M}$  or 20 – 200  $\mu\text{M}$  acetaldehyde, and 50 – 500  $\mu\text{M}$   $\text{NAD}^+$  in 50 mM sodium BES, pH 7.5 at 25°C. Reactions were initiated by adding enzyme. Data was fitted using the Cleland FORTRAN program<sup>161</sup>, while graphs were produced using SigmaPlot (StatSys v12.3), and the values represent the mean/SEM of three experiments (each  $n = 3$ ).

We determined the mode of inhibition of the compounds using steady-state kinetics by co-varying compound and substrate concentrations, with a saturating amount of the second substrate. The reactions contained 100 – 150 nM ALDH1A1 and 1% DMSO in 50 mM sodium BES at pH 7.5 at room temperature. Following a 2 minute incubation of enzyme,  $\text{NAD}^+$ , and compound, the reaction was initiated by adding aldehyde substrate and monitored at 340 nm for 3 minutes on a Beckman DU-640 spectrophotometer. To determine the mode of inhibition with respect to substrate acetaldehyde, acetaldehyde was varied from 100 – 800  $\mu\text{M}$  with a fixed concentration of  $\text{NAD}^+$  at either 800 or 1000  $\mu\text{M}$ . When cofactor  $\text{NAD}^+$  was varied, the reactions contained 20 – 250  $\mu\text{M}$   $\text{NAD}^+$  and 200  $\mu\text{M}$  propionaldehyde. All data were fit to competitive, noncompetitive,

uncompetitive, and mixed inhibition models using both single substrate-single inhibitor and tight binding inhibition programs in SigmaPlot (StatSys v12.3). The appropriate model was selected through analysis of goodness-of-fit and the residuals of those fits. The value represents the average of three independent experiments (each n = 3). Each independent experiment contained protein from a single protein prep, but at least two protein preps were used to generate the final values.

DEAB behaved as an irreversible inhibitor for ALDH1A2 and ALDH2. For ALDH1A2, solutions containing 100 nM enzyme, 0.1 – 2.0  $\mu$ M DEAB, and 1.5 mM NAD<sup>+</sup> in 50 mM sodium BES, pH 7.5 were incubated at 25°C. For ALDH2, solutions containing 100 nM enzyme, 100 – 400 nM DEAB, and 1.5 mM NAD<sup>+</sup> in 50 mM sodium BES, pH 7.5 were incubated at 25°C. At the designated time point, the remaining enzyme activity was determined spectrophotometrically by adding a saturating amount of propionaldehyde (1.0 mM) and monitoring NADH production at 340 nm on a Beckman DU-640. This concentration of propionaldehyde is sufficient to prevent any additional DEAB-dependent inactivation of both ALDH1A2 and ALDH2, thereby obviating the need for sample dilution prior to the assay. The apparent bimolecular rate constants were determined using the traditional linear analysis of covalent enzyme inactivation<sup>162</sup>:

$$\ln\left(\frac{Activity_t}{Activity_0} \times 100\right) = -k' * t$$

A secondary plot was generated from the above and the slope of this secondary plot through the origin represented the apparent bimolecular rate constant. The values represent the mean/SEM of three independent experiments, with ALDH2 at n = 3 and ALDH1A2 at n = 2.

To monitor ALDH1A1 and ALDH3A1 activity using DEAB as a substrate, it was necessary to determine both what wavelength to monitor and the differential molar extinction coefficient between DEAB and NADH at that wavelength. As the reaction time progressed, we observed a decrease in absorbance at 360 nm due to loss of DEAB and an increase in absorbance at 300 nm and 340 nm due to increases in product formation, diethylaminobenzoic acid and NADH, respectively. The largest molar

extinction coefficient difference between DEAB and NADH was seen at 360 nm, and this wavelength was used for these kinetic assays (molar extinction coefficient of  $30160 \text{ M}^{-1} \text{ cm}^{-1}$ ). Molar extinction coefficient determination and kinetics for ALDH3A1 with DEAB were performed by Dr. Bibek Parajuli.

### **7. Wavelength Scans to Monitor DEAB Oxidation**

Wavelength scans were performed on the Cary 300 Bio UV-vis spectrophotometer to monitor changes in absorbance that could indicate DEAB was acting as a substrate. Scans were performed in 1-mL quartz cuvettes and unless otherwise noted contained 500 nM enzyme, 10  $\mu\text{M}$  DEAB, and 200  $\mu\text{M}$   $\text{NAD}^+$  in 50 mM sodium BES pH 7.5 at 25°C. To confirm enzyme activity, a known substrate for each enzyme was added after completion of the scan to verify that the enzyme and assay conditions were functional. Wavelength scans for ALDH4A1 and ALDH5A1 were performed by Cameron Buchman.

### **8. Mass Spectrometry**

Complexes for analysis by mass spectrometry were formed from 2.5 – 10.0  $\mu\text{M}$  of the ALDH isoenzyme with 10  $\mu\text{M}$  DEAB with or without 100 – 500  $\mu\text{M}$   $\text{NAD}^+$  and incubated for 1 hour at room temperature in 10 mM HEPES, pH 7.5. Samples (0.5 – 5  $\mu\text{L}$ ) were injected using an Agilent 1200SL HPLC with a flow rate of 0.3 mL/min consisting of 70%  $\text{H}_2\text{O}$  and 30% acetonitrile with 0.1% formic acid into an Agilent 6520 quadrupole-time of flight (Q-TOF) mass spectrometer operating in TOF mode. The spectra were extracted and deconvoluted using MassHunter and Bioconfirm software. All mass spectrometry was performed at the Mass Spectrometry Facility, Department of Chemistry and Chemical Biology, Indiana University, Purdue University – Indianapolis with the assistance of Dr. Karl Dria.

### III. Characterization of Diethylaminobenzaldehyde

#### A. Results

##### 1. DEAB as a Substrate for ALDH Isoenzymes

DEAB shows a characteristic absorbance peak near 360 nm, while the oxidized product N,N-diethylaminobenzoic acid shows an absorbance peak at 300 nm (Figure 18). If DEAB is a substrate for members of the aldehyde dehydrogenase, it should be possible to monitor the decrease in absorbance at 360 nm and an increase in absorbance at 300 nm. The changes near these absorbance bands will be the convolution of the decrease of DEAB, increase of the acid product and the increase in NADH due to its less prominent absorption band at 340 nm ( $6220 \text{ M}^{-1} \text{ cm}^{-1}$ ). The extinction coefficient for DEAB at 360 nm is  $34380 \text{ M}^{-1} \text{ cm}^{-1}$ , and the extinction coefficient for diethylaminobenzoic acid at 300 nm is  $20370 \text{ M}^{-1} \text{ cm}^{-1}$ . In addition, there are some changes in the absorbance features over time, even in the absence of enzyme (Figure 19), possibly due to the interaction of the aldehyde with water leading to the formation of a *gem* diol. However, in the presence of five of the nine ALDH enzymes tested, there is an observable increase in the rate of oxidation, indicating that DEAB can act as a substrate for these isoenzymes. For

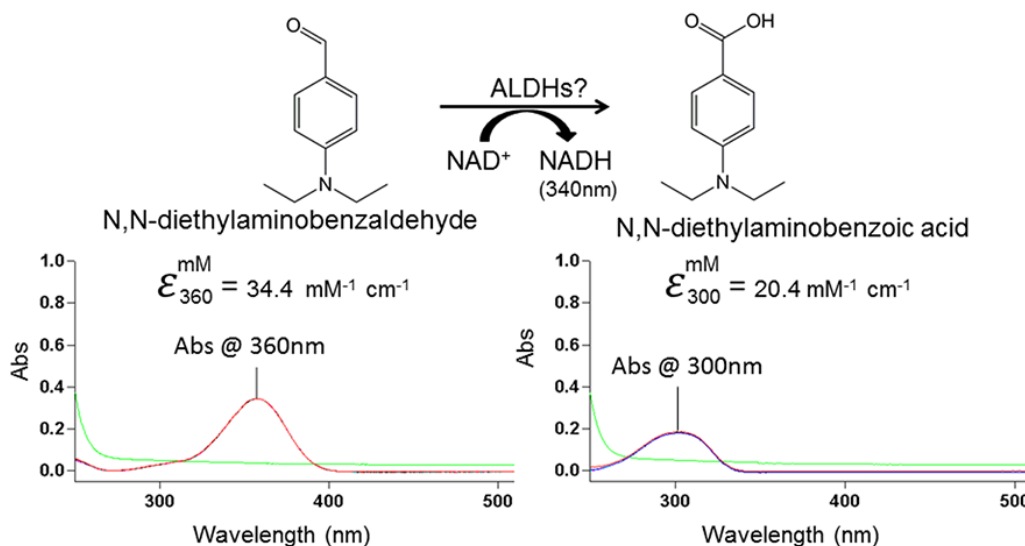


Figure 18: Absorption characteristics of DEAB and diethylaminobenzoic acid. The substrate and product of the reaction absorb at different wavelengths, allowing the reaction to be monitored. For both diethylaminobenzaldehyde and diethylaminobenzoic acid, 10  $\mu\text{M}$  of compound was used. The green trace represents the buffer BES and 100  $\mu\text{M}$  NAD<sup>+</sup>.



incubations between DEAB and ALDH1A1, the scan shows a drop in absorbance at 360 nm, indicating a loss of DEAB substrate, plus an increase in absorbance at 300 nm, corresponding to the carboxylic acid product (Figure 20A). With 500 nM ALDH1A1, oxidation of a large percentage of the 10  $\mu$ M DEAB had occurred in 3.5 hours. The strong absorption of DEAB at 360 nm partially masks NADH formation at 340 nm. Monitoring the change in absorbance at 360, we calculated the ALDH1A1 turnover rate of DEAB at  $0.028 \pm 0.002$  per min. DEAB was turned over very rapidly by ALDH3A1 using only 25 nM enzyme, indicating that it is a very good substrate for this isoenzyme (Figure 20B). However, with ALDH2 there was only a moderate change in the traces after 4 hours, similar to DEAB incubated in the absence of enzyme (Figure 20C). Wavelength scans for ALDH1A2 and ALDH4A1 were similar to ALDH2, suggesting that DEAB is either an extremely slow substrate or is not turned over by these three isoenzymes. DEAB is a very slow substrate for ALDH1B1 (Figure 20D), and the wavelength scans for ALDH1A3 and ALDH5A1 were similar to ALDH1B1. For all nine isoenzyme scans, there was no change in absorbance noted from 400 – 600 nm, and even after a 16 hour scan, the enzymes were still active and capable of propionaldehyde oxidation. Wavelength scans indicated that DEAB was turned over by ALDH3A1 much more rapidly than by the other ALDH isoenzymes tested. An analysis using substrate-inhibition kinetics produced a  $K_m$  of  $5.6 \pm 0.4$   $\mu$ M (Figure 21). For comparison, the  $K_m$  of the commonly used substrate benzaldehyde is over 30-fold higher at  $\sim 300$   $\mu$ M (Table 4).

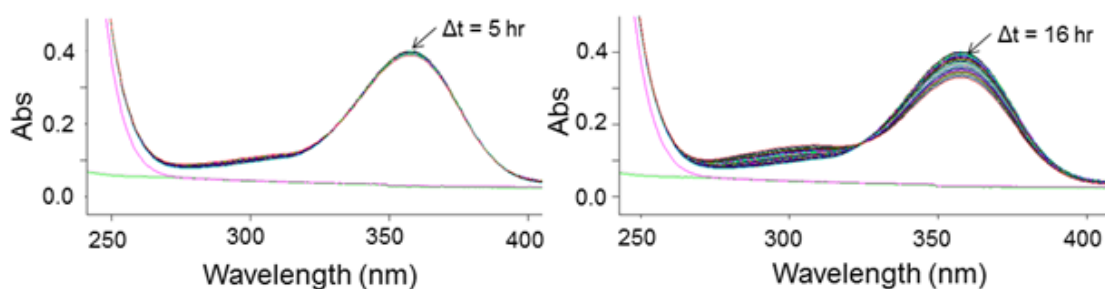


Figure 19: Oxidation of DEAB in the absence of enzyme.

Wavelength scans indicate that DEAB oxidizes to its acid at a slow rate in a 50 mM  $\text{Na}^+$ -BES pH 7.5 at 25°C, the buffer conditions used for wavelength scan assays with the enzymes. No  $\text{NAD}^+$  is present. The green line represents a water trace while the magenta line is BES alone (no DEAB). Over time, there was a decrease at 360 nm corresponding to DEAB and an increase at 300 nm corresponding to diethylaminobenzaldehyde.

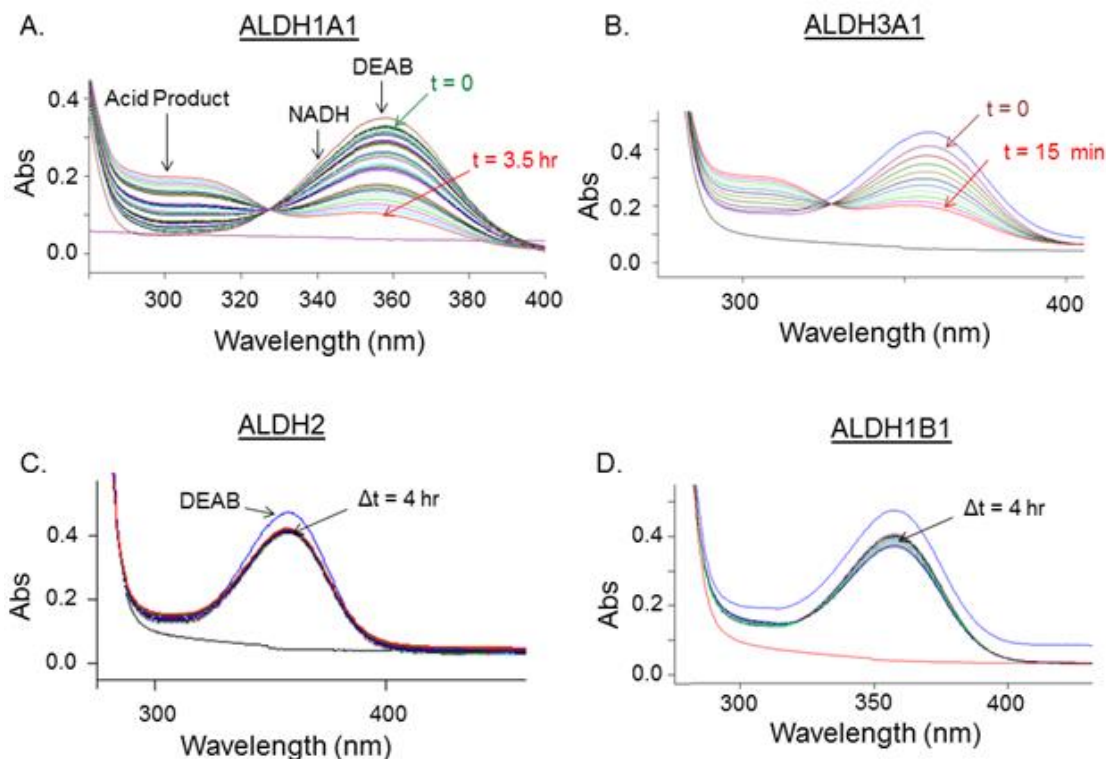


Figure 20: Wavelength scans of ALDH isoenzymes with DEAB.

For ALDH1A1, ALDH2, and ALDH1B1 reactions containing 500 nM enzyme, 200  $\mu\text{M}$   $\text{NAD}^+$ , and 10  $\mu\text{M}$  DEAB were monitored from 200 – 600 nm. A) For ALDH1A1, a gradual decrease at 360 nm and increase at 300 nm is seen, corresponding to substrate lose and product formation, respectively. B) For ALDH3A1, the reaction contained 25 nM enzyme, 200  $\mu\text{M}$   $\text{NAD}^+$ , and 10  $\mu\text{M}$  DEAB and rapid DEAB oxidation occurs. C) For ALDH2, there is little change in the spectrographic trace after 4 hours. D) ALDH1B1 had more DEAB oxidation then ALDH2 but turnover was still low. For all reactions, no absorption changes were seen above approximately 400 nm.

Table 4: Kinetic parameters for ALDH3A1 oxidation of DEAB

Substrate	$K_M$ (mM)	$k_{cat}$ ( $\text{min}^{-1}$ )	$k_{cat}/K_M$
Benzaldehyde	$280 \pm 20$	$1400 \pm 60$	$4.9 \pm 0.3$
DEAB	$5.6 \pm 0.7$	$130 \pm 20$	$24 \pm 1$

Bibek Parajuli, Ph.D.

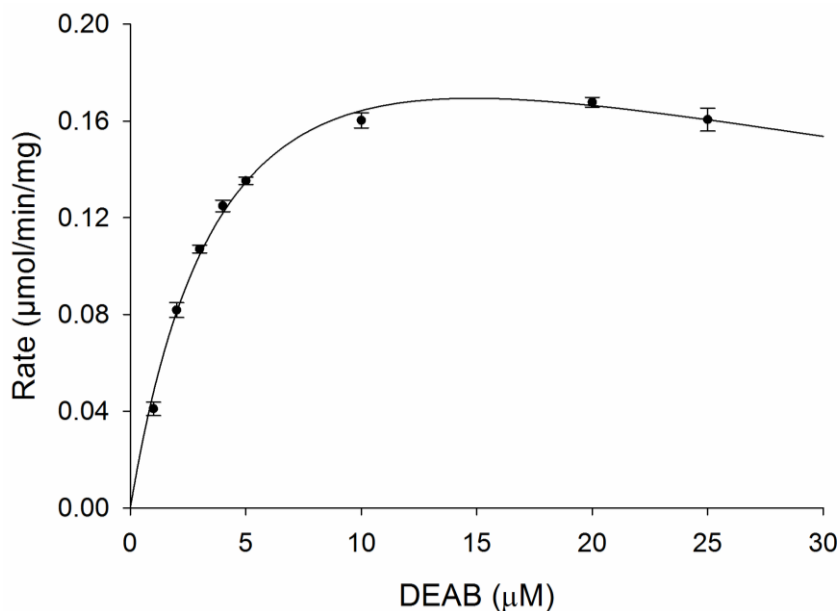


Figure 21: Substrate saturation curve for ALDH3A1 with DEAB as a substrate. This curve represents one of three experiments, each  $n = 3$ , fit to the substrate inhibition equation (Bibek Parajuli, Ph.D.).

## 2. DEAB as an Inhibitor of ALDH Isoenzymes.

Although DEAB is a substrate for ALDH1A1, ALDH1A3, ALDH1B1, and ALDH5A1, it is turned over at such a slow rate it is effectively an inhibitor for the oxidation of other substrate aldehydes such as acetaldehyde and propionaldehyde. DEAB inhibited six ALDH isoenzymes tested, and all had  $IC_{50}$  values less than  $20 \mu\text{M}$  (Figure 22 and Table 5). Under the conditions of the experiments, it was a more potent inhibitor for ALDH1A1 (Figure 23A,  $IC_{50} = 57 \pm 5 \text{ nM}$ ) compared to the other five isoenzymes inhibited by DEAB. DEAB displayed a competitive mode of inhibition toward ALDH1A1 with respect to varied acetaldehyde with a  $K_i$  of  $9.8 \pm 3.1 \text{ nM}$  (Figure 23B). During  $IC_{50}$  calculations for ALDH2, it was discovered that the  $IC_{50}$  values were dependent on order of addition and incubation times, with no inhibition observed at up to  $50 \mu\text{M}$  (maximum concentration used) when the reaction was initiated by adding enzyme, but dropped to an  $IC_{50}$  of  $160 \pm 30 \text{ nM}$  following a two minute pre-incubation of enzyme with DEAB and  $\text{NAD}^+$  (Figure 24). Under these conditions, saturation with propionaldehyde did not lead to recovery of full enzymatic activity, suggesting that DEAB is covalently modifying ALDH2. Although not as apparent as with ALDH2, ALDH1A2 behaved in a similar manner, with the  $IC_{50}$  for ALDH1A2 dependent on the order of addition and exhibited

more than a 5-fold increase without a pre-incubation with DEAB and  $\text{NAD}^+$ . The apparent bimolecular rate constant of ALDH1A2 was  $2900 \pm 160 \text{ M}^{-1}\text{s}^{-1}$  (Figure 25), while ALDH2 had an apparent bimolecular rate constant of  $86000 \pm 4200 \text{ M}^{-1}\text{s}^{-1}$  (Figure 26).

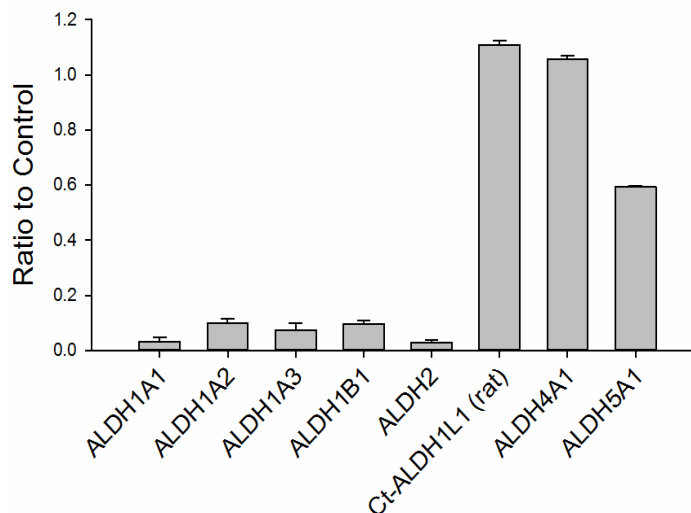


Figure 22: Selectivity of DEAB with eight ALDH isoenzymes. For each enzyme, activity following treatment with  $20 \mu\text{M}$  DEAB was compared to control activity with 1% DMSO. Bar represents mean/SEM ( $n=3$ ).

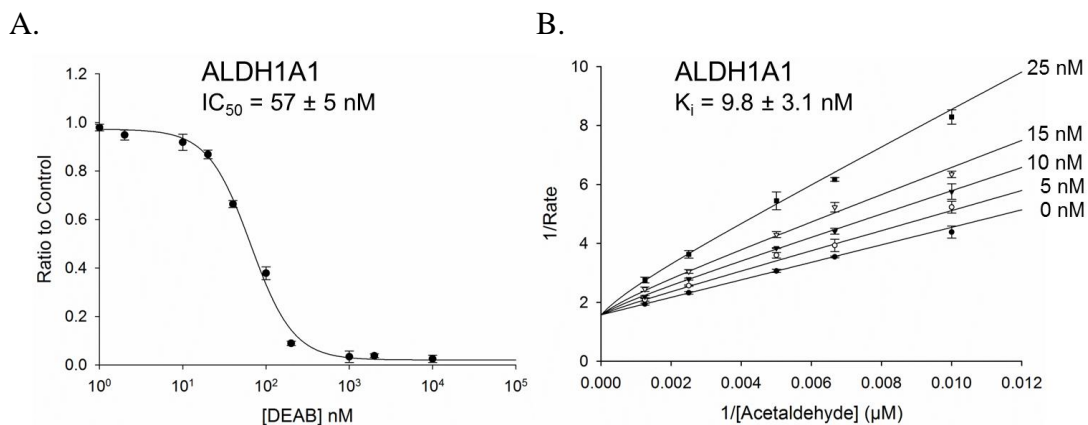


Figure 23: Characterization of ALDH1A1 with DEAB.

A)  $\text{IC}_{50}$  curve of DEAB with  $150\text{nM}$  ALDH1A1. B) Lineweaver-Burk plot to determine the  $K_i$  of DEAB for ALDH1A1 with respect to varied acetaldehyde. Each plot represents one of three experiments, with each point representing the mean/SEM of three independent readings.

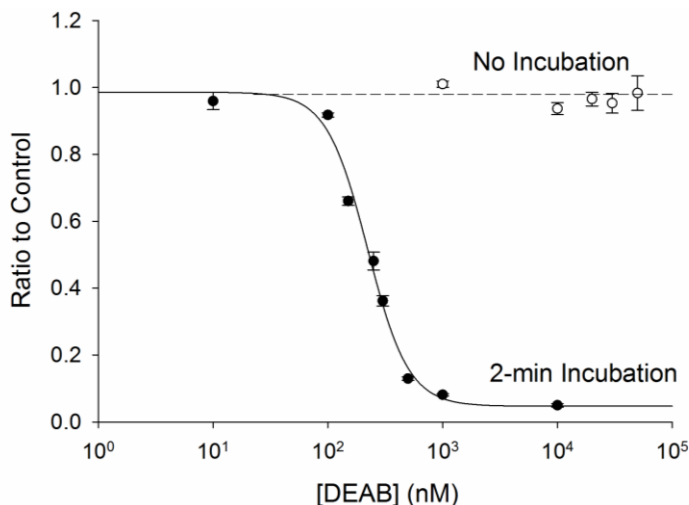


Figure 24: ALDH2 and DEAB.

With a 2 min incubation of ALDH2 with DEAB and  $\text{NAD}^+$ , DEAB is a potent inhibitor with an  $\text{IC}_{50} = 160 \text{ nM}$ . Without an incubation period, no inhibition was seen at up to  $50 \mu\text{M}$  DEAB.

Table 5: Effect of DEAB on ALDH isoenzymes.

	Substrate	$\text{IC}_{50} (\mu\text{M})$ Mean/SEM	Comments
ALDH1A1	Yes	$0.057 \pm 0.005$	Competitive Tight Inhibition $K_i = 10 \pm 3 \text{ nM}$
ALDH1A2	No	$1.2 \pm 0.1$	Covalent Inhibitor Apparent Bimolecular Rate Constant $2900 \pm 160 \text{ M}^{-1} \text{ s}^{-1}$
ALDH1A3	Yes (slow)	$3.0 \pm 0.3$	
ALDH1B1	Yes (slow)	$1.2 \pm 0.1$	
ALDH1L1	No	NI*	
ALDH2	No	$0.16 \pm 0.03$	Covalent Inhibitor Apparent Bimolecular Rate Constant $86000 \pm 4200 \text{ M}^{-1} \text{ s}^{-1}$
ALDH3A1	Yes	NA#	$K_m = 5.6 \mu\text{M}$
ALDH4A1	No	NI*	
ALDH5A1	Yes (slow)	$13 \pm 0.5$	

\*NI denotes no detectable inhibition. #NA denotes not applicable as an inhibitor. ALDH3A1  $K_m$  determined by Bibek Parajuli, Ph.D. ALDH4A1 and ALDH5A1 substrate analysis and ALDH5A1  $\text{IC}_{50}$  performed by Cameron Buchman.

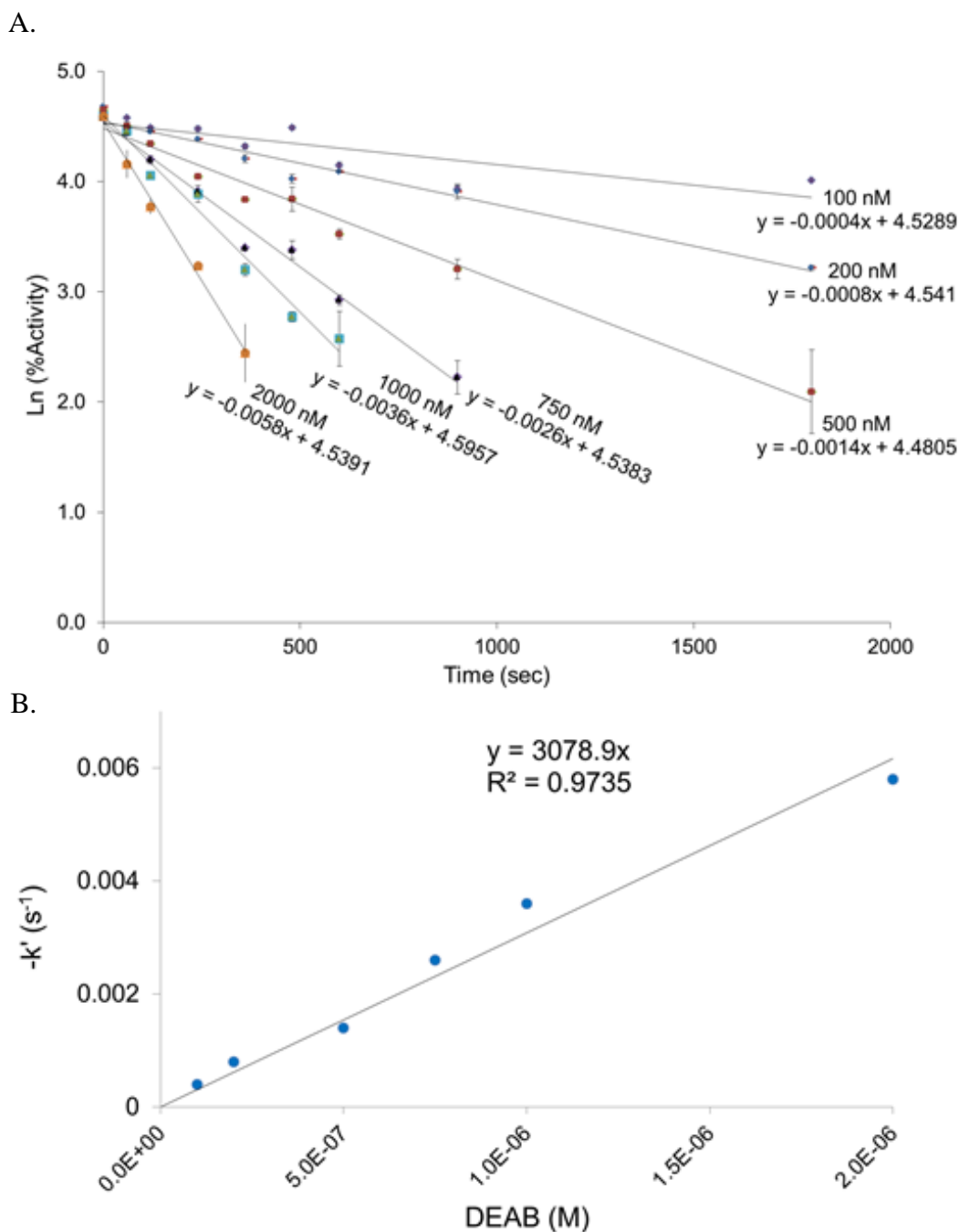


Figure 25: Time-dependent inhibition of human ALDH1A2 with DEAB. A) Plot of  $\ln$  (% activity) vs. time with 6 doses of DEAB. The slopes ( $-m$ ) of these lines were used to generate the secondary plot. B) Secondary plot of  $-k'$  versus DEAB concentration for ALDH1A2. The bimolecular rate constant for inhibition was calculated from the slope of this line. Each set of plots represents one of three experiments, each  $n = 2$ .

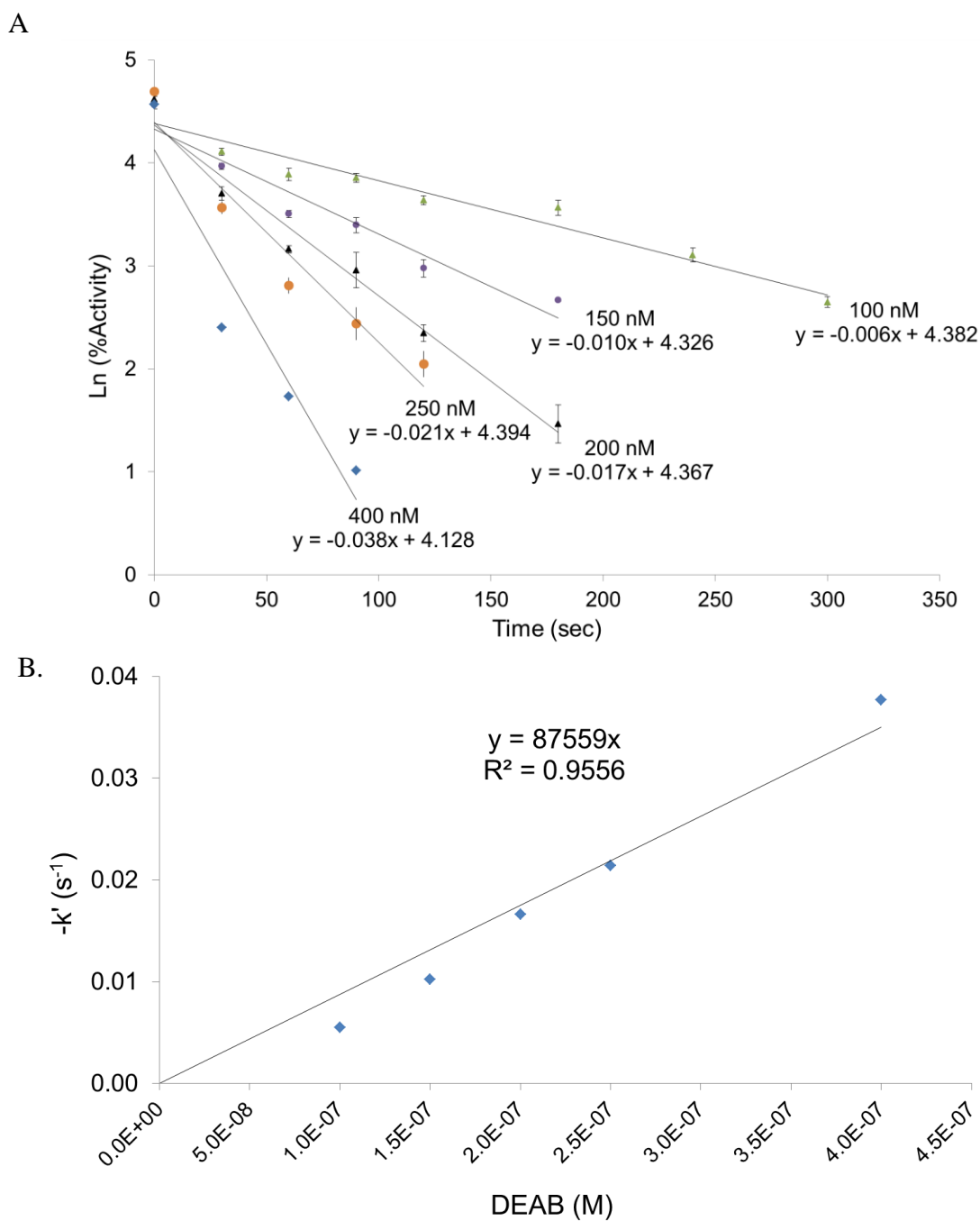


Figure 26: Time-dependent inhibition of human ALDH2 with DEAB.

A) Plot of  $\ln$  (% activity) vs. time with 5 doses of DEAB. The slopes ( $-m$ ) of these lines were used to generate the secondary plot. B) Secondary plot of  $-k'$  versus DEAB concentration for ALDH2. The bimolecular rate constant for inhibition was calculated from the slope of this line. Each set of plots represents one of three experiments, each  $n = 3$ .

### 3. Mass Spectrometry of ALDH-DEAB Complexes.

Mass spectrometry was used to determine whether DEAB was covalently binding to the ALDH enzymes. Quadrupole TOF MS of ALDH1A1, ALDH1A2, ALDH1B1, and ALDH2 confirmed that there was an increase in mass of 175 Da, corresponding to DEAB binding with the loss of two protons. With 5  $\mu\text{M}$  ALDH2 alone (Figure 27A) or incubated with 10  $\mu\text{M}$  DEAB without  $\text{NAD}^+$  (Figure 27B), there was a major peak at a mass of 54445 Da plus a much smaller peak at 54357 Da, a difference of 87 Da likely corresponding to loss of the N-terminal serine during protein production. ALDH2 incubated with both DEAB and 500  $\mu\text{M}$   $\text{NAD}^+$  resulted in a major peak at 54620 Da and much smaller peaks at 54533 Da, 54445 Da, and 54357 Da (Figure 27C). The peaks at 54620 Da and 54533 Da represent a shift of 176 Da from 54445 Da and 54357 Da, respectively, and correspond to DEAB binding to the two major ALDH2 species present in the sample. A similar shift was seen using 2.5  $\mu\text{M}$  ALDH1B1, 10  $\mu\text{M}$  DEAB, and 500  $\mu\text{M}$   $\text{NAD}^+$  (data not shown). ALDH1B1 alone or incubated with DEAB yielded a major

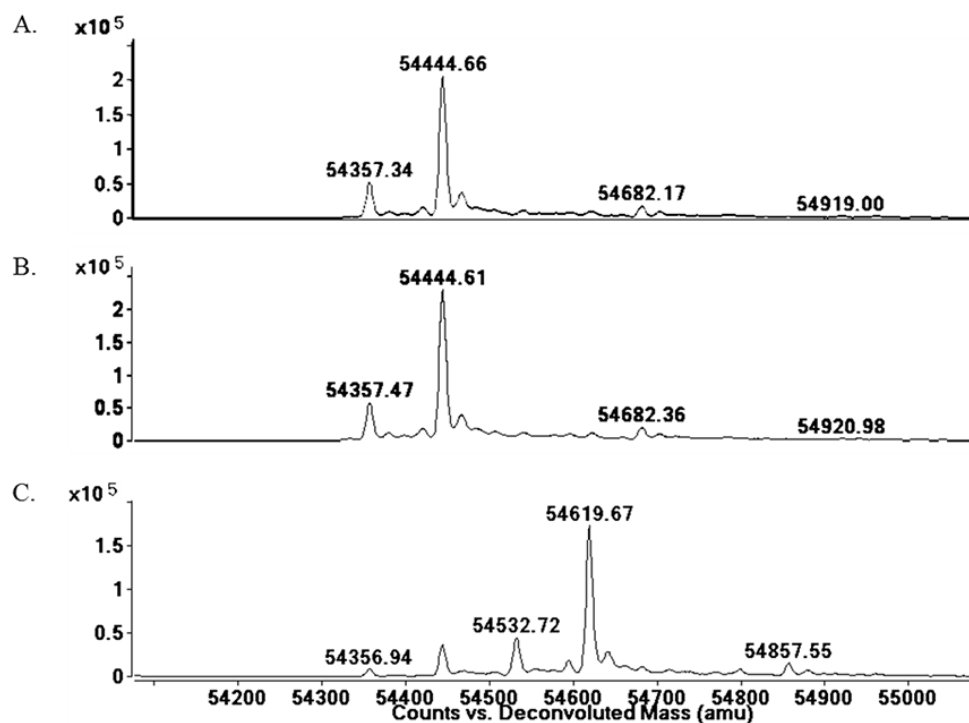


Figure 27: Modification of human ALDH2 with DEAB. Deconvoluted spectrum of A) ALDH2, B) ALDH2 incubated with 2-fold molar excess of DEAB, and C) ALDH2 incubated with 2-fold molar excess of DEAB and saturating  $\text{NAD}^+$ .



peak at 56592 Da and a much smaller peak at 57200 Da; this latter peak may represent an enzyme species modified by two glutathione molecules. However, incubation of ALDH1B1, DEAB, and  $\text{NAD}^+$  shifted the major peak to 56768 Da plus a minor peak at 57375 Da, both corresponding to DEAB bound to the two ALDH1B1 species present, and a smaller peak at 56592 Da, corresponding to unmodified protein. Both the ALDH2 and ALDH1B1 data indicate that  $\text{NAD}^+$  must be present for DEAB to modify the enzyme. After a one hour incubation with a 2- to 4-fold molar excess of DEAB, most of the enzyme present had been modified. DEAB also covalently modifies ALDH1A1 in the presence of  $\text{NAD}^+$  but is not complete (Figure 28). These shifts in peaks correspond to DEAB binding with a loss of two protons, similar to what is seen with ALDH1A1 and a known substrate benzaldehyde. Using 5  $\mu\text{M}$  ALDH1A1, 1.5 mM  $\text{NAD}^+$ , and 2.5 mM benzaldehyde, there is a 104 Da shift corresponding to the substrate binding to ALDH1A1 with the loss of two protons in the presence of enzyme, cofactor, and substrate (Figure 29). Association of DEAB to ALDH1A1, ALDH1B1, and ALDH2 was

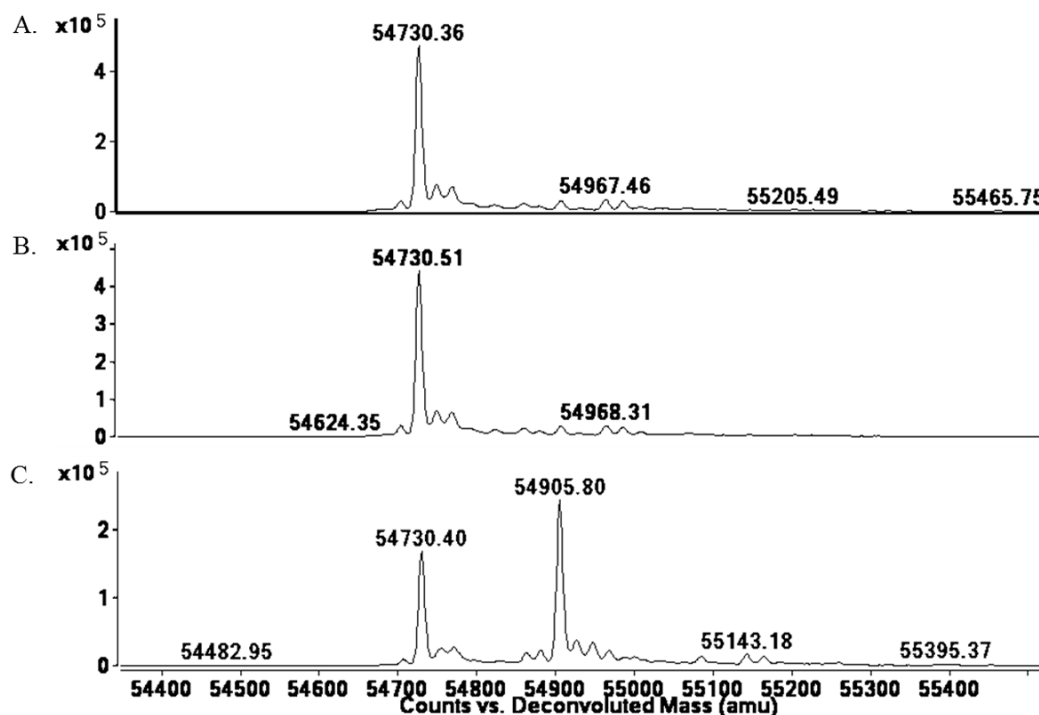


Figure 28: Modification of human ALDH1A1 with DEAB. Deconvoluted spectrum of A) ALDH1A1, B) ALDH1A1 incubated with 10-fold molar excess of DEAB, and C) ALDH1A1 incubated with 10-fold molar excess of DEAB and saturating  $\text{NAD}^+$ .

dependent on  $\text{NAD}^+$ , but the requirement for  $\text{NAD}^+$  was not absolute for ALDH1A2. Using  $5.0 \mu\text{M}$  ALDH1A2,  $10 \mu\text{M}$  DEAB, and  $500 \mu\text{M}$   $\text{NAD}^+$ , enzyme alone yielded a major peak at  $56592 \text{ Da}$  and a much smaller peak at  $56830 \text{ Da}$  (Figure 30A). When ALDH1A2 is incubated with DEAB, the major peak is still at  $56592 \text{ Da}$  but a large peak is also seen at  $56768 \text{ Da}$ , corresponding to DEAB binding to approximately one-third of the ALDH1A2 present, with smaller peaks at  $56830 \text{ Da}$  and  $57007 \text{ Da}$  (Figure 30B). Following incubation of ALDH1A2, DEAB, and  $\text{NAD}^+$ , the major peak was seen at  $56768 \text{ Da}$ , with smaller peaks at  $56592 \text{ Da}$  and  $57006 \text{ Da}$  (Figure 30C). These results imply that DEAB may bind to ALDH1A2 independent of the presence of coenzyme. However, it may be possible that oxidized coenzyme co-purifies with ALDH1A2. To determine whether  $\text{NAD}^+$  is present from the ALDH1A2 protein purification,  $500 \mu\text{M}$  propionaldehyde was added to  $18 \mu\text{M}$  ALDH1A2 without additional  $\text{NAD}^+$  and monitored for 15 minutes. There was no change at  $340 \text{ nm}$ , indicating that  $\text{NADH}$  is not being produced in this reaction. This suggests that  $\text{NAD}^+$  is, at a minimum, not present at the molar ratio levels suggested by mass spectrometry, since reactions using only  $5 \mu\text{M}$   $\text{NAD}^+$  produce an observable change at  $340 \text{ nm}$ . To confirm covalent modification and

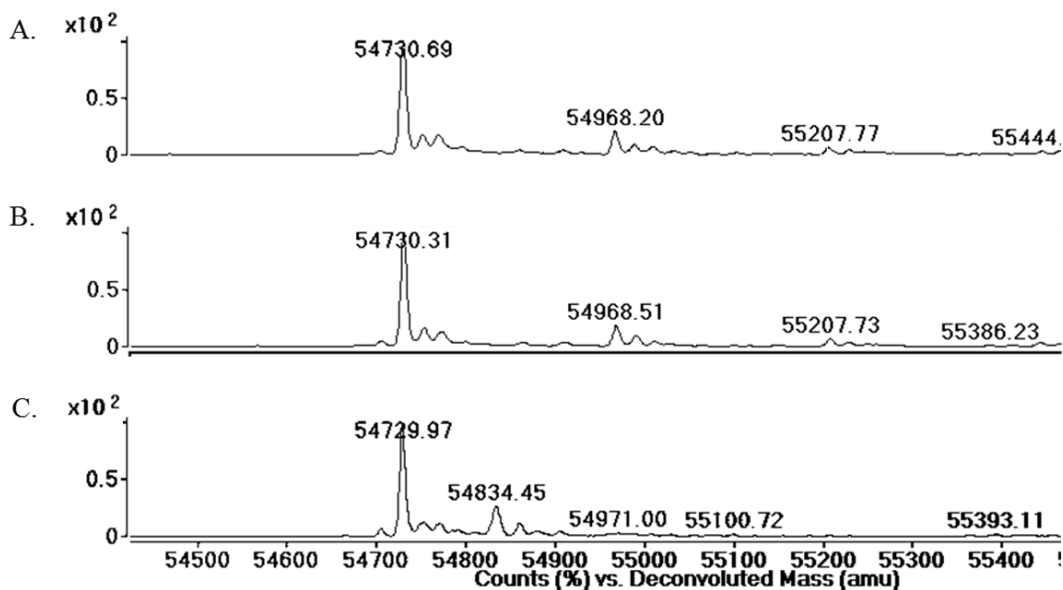


Figure 29: Modification of human ALDH1A1 with benzaldehyde. Deconvoluted spectrum of A) ALDH1A1, B) ALDH1A1 incubated with 500-fold molar excess of benzaldehyde, and C) ALDH1A1 incubated with 500-fold molar excess of benzaldehyde and saturating  $\text{NAD}^+$ .

investigate the requirement for  $\text{NAD}^+$  to inactivate ALDH2 and ALDH1A2, the enzymes were incubated in the presence of DEAB with and without  $\text{NAD}^+$  and then dialyzed for 4 hours against a 4000-fold excess of buffer. Relative to control samples incubated in the absence of DEAB, neither the ALDH2 nor the ALDH1A2 samples regained enzymatic activity when DEAB and  $\text{NAD}^+$  were present in the incubation, consistent with irreversible, covalent modification (Figure 31). In contrast to the level of inactivation observed with DEAB and  $\text{NAD}^+$ , incubation of enzyme plus DEAB alone, resulted in only partial inactivation of both ALDH2 and ALDH1A2. Dialysis had no impact on the level of inactivation for ALDH2, but the ALDH1A2 sample recovered 75% of the control activity, suggesting the latter modification is, in part, reversible (Figure 31). ALDH1L1, ALDH3A1, ALDH4A1 and ALDH5A1 were not examined in this manner due to limited quantities of enzyme available.

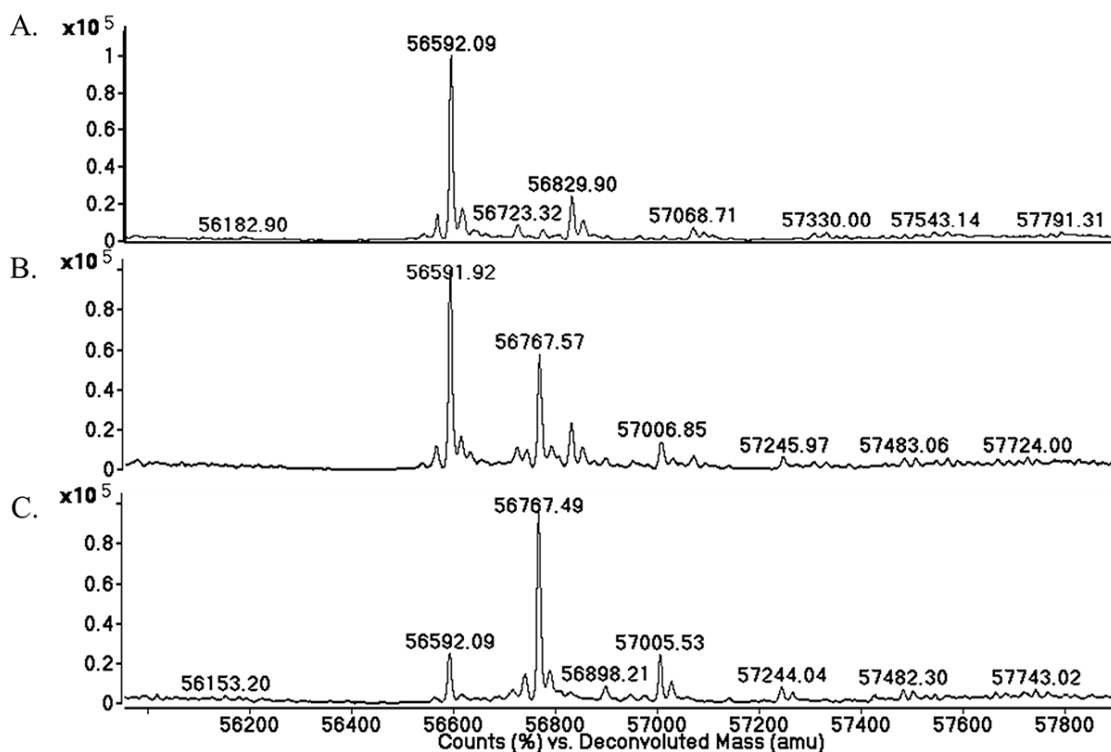


Figure 30: Modification of human ALDH1A2 with DEAB. Deconvoluted spectrum of A) ALDH1A2, B) ALDH1A2 incubated with 2-fold molar excess of DEAB, and C) ALDH1A2 incubated with 2-fold molar excess of DEAB and saturating  $\text{NAD}^+$ .

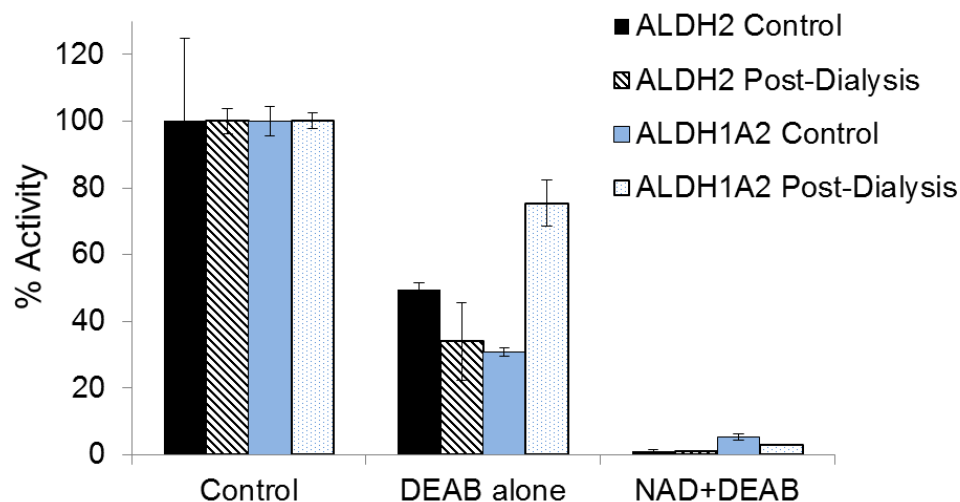


Figure 31: Effects of dialysis on DEAB-induced inhibition.

ALDH1A2 and ALDH2 were pre-incubated for 30 minutes with DEAB in the presence or absence of  $\text{NAD}^+$ . Solutions were dialyzed for 4 hours against a 4000-fold excess of buffer and the activity post-dialysis compared to the activity levels prior to dialysis. Bars represent mean/SEM (n=3) (performed by Cameron Buchman).

## B. Discussion

A number of isoenzymes in the aldehyde dehydrogenase superfamily have been linked to both normal and cancer stem cells. The Aldefluor Assay relies on their aldehyde oxidation abilities to identify and segregate stem cells with high levels of ALDH activity. Inhibition of ALDHs by DEAB, an aldehyde, is used as a control. The concept of substrates as inhibitors is not new. A number of organophosphorus compounds are esters that inhibit cholinesterase by acetylating the enzyme in a process similar to its normal substrate acetylcholine. As discussed by Aldridge, the only difference between a substrate (i.e. acetylcholine) and an inhibitor (i.e. dimethylphosphates) is the rate of reaction<sup>163</sup>. Here, we have shown that DEAB is both a substrate and an inhibitor for certain ALDH isoenzymes and that the rate of reaction, namely hydrolysis, determines whether the compound behaves as a classic substrate (ALDH3A1), as a covalent inhibitor, (ALDH2 and ALDH1A2), or as an intermediate between substrate and inhibitor (ALDH1A1, ALDH1A3, ALDH1B1, and ALDH5A1). We propose that differences in residues at the substrate binding sites of the isoenzymes are capable of stabilizing the acyl-enzyme intermediate to varying degrees. In ALDH3A1, little to no stabilization

occurs and DEAB is quickly turned over, resulting in a low micromolar  $K_m$ . For ALDH2 and ALDH1A2, the acyl-enzyme is very stable and little observable turnover occurs, thus mimicking a covalent inhibitor. However, we do believe these two enzymes are capable of eventually oxidizing DEAB at an extremely low rate. We propose that structural features within their respective active sites are capable of stabilizing a resonance structure intermediate (Figure 32).

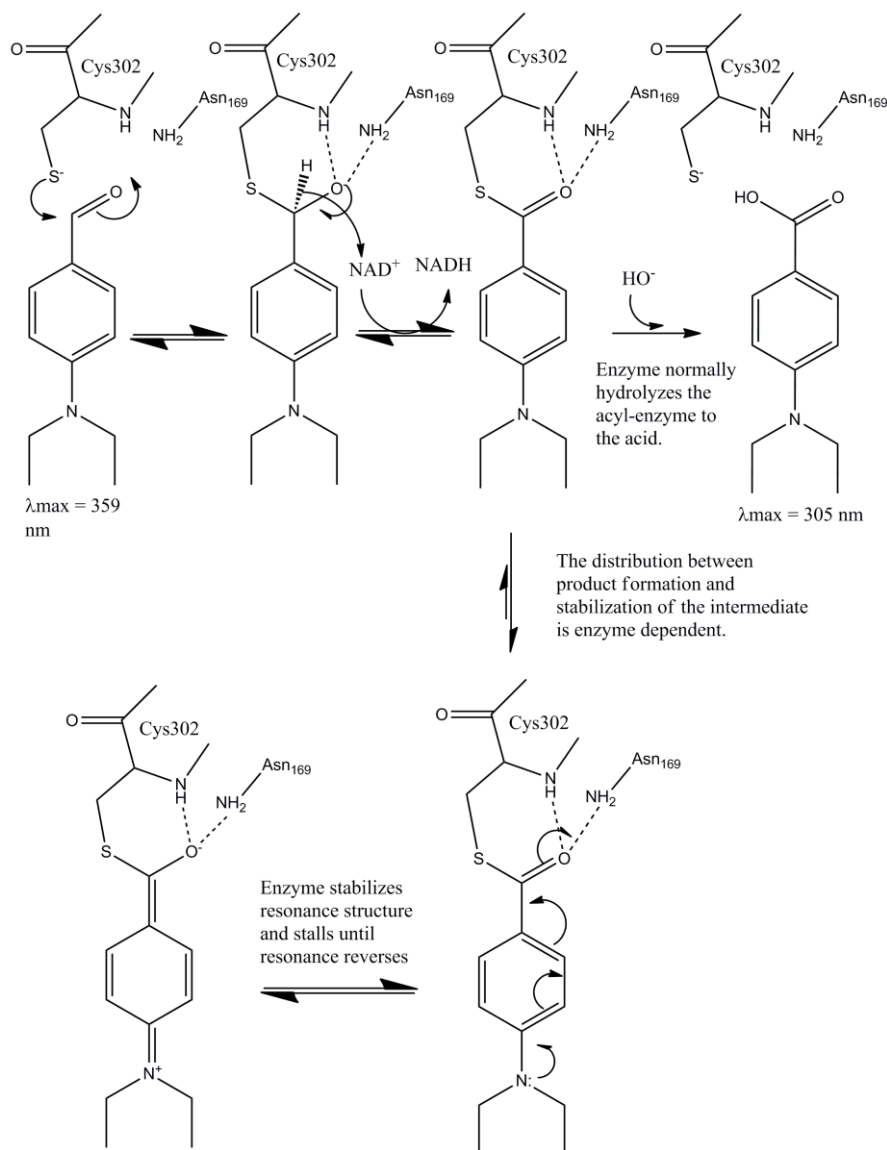


Figure 32: Proposed mechanism of action of DEAB on ALDH enzymes. Residue numbering based on ALDH2.

Based on the kinetic and mass spectrometry data, we propose that DEAB functions as a substrate for all enzymes through the hydride transfer step, but that the acyl-enzyme intermediate is slowed for hydrolysis by a resonance structure available to the thioester intermediate due to the lone pair of electrons on the para-diethylamino substituent (Figure 32). The addition of the lone pair into the aromatic ring resonance promotes the formation of a quinoid-like structure where the extra electrons now reside on the carbonyl oxygen atom resulting in stabilization by the oxy-anion hole comprised of the peptide amide nitrogen of the catalytic nucleophile and the side chain amide nitrogen of the residue equivalent to Asn169 in ALDH2. The intermediate stabilized in this manner makes the carbonyl carbon less susceptible to attack by a hydroxyl ion. The extent that this resonance structure forms and the duration of its stabilization would appear to differ between the various isoenzymes with which DEAB forms productive interactions.

The fact that DEAB is an excellent substrate for ALDH3A1 is not altogether surprising, since the enzyme is known to prefer aromatic substrates over linear aliphatic substrates<sup>164,165</sup>. It is interesting that even for ALDH3A1, the  $k_{cat}$  is lower for DEAB than for benzaldehyde by an order of magnitude, making DEAB, by definition, a slow substrate. Thus, only the fact that the  $K_m$  for DEAB is approximately 50-fold lower than for benzaldehyde places the catalytic efficiency for DEAB amongst the best for known substrates. For acetaldehyde and benzaldehyde, hydride transfer is known to be rate-limiting for catalysis in ALDH3A1<sup>42</sup>, implying that either hydride transfer is 11-fold slower for DEAB or another step has become rate-limiting. We suggest that acyl-enzyme hydrolysis has become rate-limiting due to the potential for resonance stabilization of the acyl-enzyme intermediate and the fact that the electronic properties of DEAB should promote easier extraction of the hydride ion from the thiohemiacetal. Measurements of the isotope effects or stopped-flow analyses will be necessary to distinguish between these or other mechanistic possibilities for ALDH3A1. However, it is clear that the acyl-enzyme intermediate formed between ALDH3A1 and DEAB is the shortest lived of any ALDH acyl-enzyme species tested by at least 3-orders of magnitude. The short life-time may be influenced by the curved nature of the enzyme's substrate-binding site<sup>166</sup>, which

likely distorts the planar conformations required to maintain the aromatic quinoid resonance state (see Figure 12 for comparisons of ALDH substrate-binding sites).

DEAB is a good to excellent inhibitor for all but three of the ALDH isoenzymes examined, with  $IC_{50}$  values ranging between 57 nM to 13  $\mu$ M. The most potent inhibition by DEAB is toward ALDH1A1, followed by ALDH2, ALDH1A2, ALDH1B1, ALDH1A3 and ALDH5A1. However, the relative order for enzymatic turnover follows the order ALDH1A1 > ALDH1B1, ALDH1A3, ALDH5A1 > ALDH1A2, ALDH2. Thus, it would seem that the two properties of DEAB, inhibition vs substrate, are not directly correlated. DEAB demonstrates the most potent inhibition for ALDH1A1 with an  $IC_{50}$  of 57 nM and exhibits a competitive mode of inhibition toward acetaldehyde with a  $K_i$  of 10 nM. Consequently, despite the low turnover of this complex ( $0.028 \text{ min}^{-1}$ ), these properties give rise to a pseudo-bimolecular rate constant of approximately  $47,000 \text{ M}^{-1}\text{s}^{-1}$ , which is within a factor of two of the value obtained for DEAB inactivation of ALDH2 at  $86,000 \text{ M}^{-1}\text{s}^{-1}$ . Both these values imply a rapid association between enzyme and DEAB. ALDH1A2 is 20- to 30-fold slower, correlating with its 10- to 100-fold higher  $IC_{50}$  value. The interactions between DEAB and ALDH1A3 or ALDH1B1 are similar to ALDH1A2, although these enzymes do show very slow turnover of their respective enzyme-DEAB complexes.

An inspection of the respective active site structures fails to identify consistent differences in residues lining the substrate sites that correlate with all the features of DEAB interactions<sup>167 166</sup>. ALDH1A1 is the most wide open substrate binding site which is best described as a flattened funnel shape, due to Gly substitutions at positions 125 and 458 and Val at positions 174 and 460<sup>166</sup>. The ALDH1A2, ALDH1A3 and ALDH1B1 enzymes have substrate sites more similar to the cylindrical site in ALDH2<sup>166</sup>, and perhaps these sites better constrain the aromatic system to remain planar, thereby increasing the residency time of the quinoid intermediate and slowing hydrolysis. The active site structure of ALDH5A1 (PDB code 2W8R<sup>168</sup>) is next most similar to ALDH2, but its low reactivity toward DEAB may be due to the presence of two characteristic arginine residues near the middle and outer regions of the substrate site, R213 and R334.

Presumably, the binding site in ALDH5A1 does not stabilize the acyl-enzyme intermediate sufficiently to prevent turnover. The failure of DEAB to inhibit ALDH1L1 and ALDH4A1 suggests that elements of the substrate binding site in these isoenzymes are incompatible with formation of a productive thiohemiacetal intermediate with DEAB in the first place. ALDH1L1 has a very narrow active site surrounding the catalytic nucleophile, which may exclude aromatic substrates from its active site. In agreement with this observation, aromatic aldehydes are poor substrates for ALDH1L1. In the case of ALDH4A1 (PDB code 4LH2<sup>169</sup>), it is the outer region of the substrate binding site that is narrowed, which may prevent DEAB association.

DEAB possesses all the characteristics of a substrate for many ALDH isoenzymes, especially ALDH3A1. However, its particular electronic properties give rise to a stalled acyl-enzyme intermediate in ALDH isoenzymes that can stabilize its electronic resonance structure, making DEAB a mechanism-based inhibitor for ALDH2 and ALDH1A2, a very slow substrate for ALDH1A3, ALDH1B1 and ALDH5A1, a slow substrate for ALDH1A1, and an excellent substrate for ALDH3A1. In contrast, DEAB is neither a substrate nor an inhibitor for ALDH1L1 or ALDH4A1. The broad and varied nature of the interaction between DEAB and ALDH isoenzymes suggests that its use as an inhibitor or competitive substrate in the Aldefluor Assay should be interpreted with caution with respect to which particular ALDH isoenzymes contribute to the observed signal in the flow-cytometry assay.



## IV. Discovery of Novel, Selective Inhibitors of ALDH1A1

### A. Results

#### 1. X-ray Crystallography of apo-ALDH1A1 and ALDH1A1-NADH Complex

To better understand target specificity and provide a structural context for the mechanism underlying how a compound inhibits ALDH1A1, it was necessary to determine the structure of ALDH1A1 using X-ray crystallography. To accomplish this objective, we first needed to crystallize human ALDH1A1. The structure of sheep ALDH1A1 was determined roughly 15 years ago<sup>148,149</sup>. Although the sheep and human enzymes share approximately 90% sequence identity, the conditions used to crystallize the sheep enzyme did not successfully crystallize the human enzyme. To determine crystallization conditions for human ALDH1A1, a number of Hampton Research screens were used. These screens contain a diverse range of reagents to facilitate discovery of conditions in which the protein begins to crystallize, rather than simply staying in solution or precipitating out. Only the Index Screen (HR2-144) generated positive results, with five conditions producing crystalline needle clusters (Figure 33A): (1.) 200 mM sodium chloride, 100 mM BisTris pH 6.5, 25% PEG3350, (2.) 200 mM sodium chloride, 100 mM HEPES pH 7.5, 25% PEG3350, (3.) 200 mM sodium chloride, 100 mM Tris pH 8.5, 25% PEG3350, (4.) 200 mM ammonium acetate, 100 mM BisTris pH 6.5, 25% PEG3350, and (5.) 200 mM sodium formate, 20% PEG3350. The first four conditions were successfully replicated, and these conditions were further optimized to produce multiple, small needle clusters. Next, the Additive Screen (HR2-428) was used with each of the four conditions above to identify small molecule(s) that might aid in the formation



Figure 33: Evolution of ALDH1A1 crystallization conditions.

A) Needle cluster of crystalline ALDH1A1 enzyme produced from five conditions in Hampton Research Index Screen. B) Small crystals that formed in presence of yttrium chloride. C) Final crystal produced following optimization with ytterbium chloride.

of crystals suitable for X-ray structural analysis. The salt yttrium chloride produced small crystals in the two conditions that used the buffer BisTris, PEG3350, and either sodium chloride or ammonium acetate (Figure 33B). However, further optimization with yttrium chloride did not improve crystal formation. Yttrium is a trivalent ion with an ionic radius of 0.9 Å. Other ions of similar size were tested, including gold (III) chloride, palladium (II) chloride, thulium (III) chloride, lutetium (III) chloride, ytterbium (III) chloride, and samarium (III) chloride. Ytterbium (III) chloride (YbCl<sub>3</sub>) produced large, single crystals exhibiting strong X-ray diffraction; it is a trivalent ion and with a radius of 0.858 Å, it is only slightly smaller than yttrium. YbCl<sub>3</sub> with ammonium acetate produced large, cube-shaped crystals, while YbCl<sub>3</sub> with sodium chloride produced tetragonal crystals, with the unequal edge at least four times longer than the lengths of the other two sides (Figure 33C). The cube-shaped crystals were more difficult to produce and generated X-ray diffraction data indicating a more disordered crystal. Therefore, the conditions that produced tetragonal-shaped crystals were deemed the best for X-ray crystallography structure determination. For human ALDH1A1, these conditions are 100 mM sodium BisTris, pH 6.2 - 7.5, 7 - 12 % PEG3350, 200 mM NaCl, 5 - 10 mM YbCl<sub>3</sub>, and 3 - 5 mg/mL enzyme.

A similar approach was taken to determine crystallization conditions for the co-crystals of ALDH1A1 with its cofactor NAD<sup>+</sup>. The Index Screen identified similar conditions as found with enzyme alone. Follow-up with the Additive Screen indicated nearly half of the 96 additives aided crystal formation. However, further optimization of numerous conditions did not improve crystal size, and the crystals produced did not diffract as expected for a protein. To determine the structure of ALDH1A1 bound to its cofactor, apo-enzyme crystals were soaked for two hours with 1 mM NAD<sup>+</sup>.

The structure of human ALDH1A1 had not been previously reported, but was expected to be similar to both the human ALDH2 enzyme (PDB code 3N80), with which ALDH1A1 shares about 70% sequence identity, and the sheep ALDH1A1 (PDB code 1BXS), with over 90% sequence identity. We solved the structure of human ALDH1A1 N121S in its apo-enzyme form to a resolution of 1.75 Å (Figure 34, Table 6, PDB Code 4WJ9). The

enzyme's structure is similar to other ALDH enzymes and contains all the structural features that characterize the ALDH superfamily. The structure of wild-type ALDH1A1 with NADH was determined to a resolution of 2.1 Å (Figure 35, Table 6, PDB Code 4WB9). Comparison of the structure of ALDH1A1 (PDB Code 4WB9), ALDH2 (PDB Code 10O2), and ALDH3A1 (PDB Code 4L2O) with cofactor illustrates the difficulty of developing selective inhibitors for ALDH1A1 that target the cofactor binding site. Of the three isoenzymes compared, ALDH3A1 is the least similar both by structural topology and sequence identity, as is expected based on its ability to utilize both NAD<sup>+</sup> and NADP<sup>+</sup>, and these differences are most obvious near the adenosine ribose and pyrophosphate binding site. However, as shown in Figure 36, there is a high degree of similarity between the cofactor binding sites of ALDH1A1 and ALDH2, supporting our hypothesis that selection of an assay independent of the cofactor binding site would aid development of selective inhibitors for the ALDH1/2 class of enzymes.

Table 6: Data collection and refinement statistics.

<b>Data Collection</b>	<b>Apo-ALDH1A1 N121S (PDB: 4WJ9)</b>	<b>ALDH1A1-NADH WT (PDB: 4WB9)</b>
Space Group	P422	P422
Cell Dimensions		
a, b, c (Å)	109, 109, 83	109, 109, 83
$\alpha, \beta, \gamma$ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	50 – 1.75	50 – 2.1
R <sub>merge</sub>	0.056(0.59)	0.09(0.52)
I/ $\sigma$ <sub>i</sub>	31.5 (3.9)	17.8 (4.9)
Completeness (%)	99 (97)	99 (100)
Redundancy	9.6 (8.5)	8.3 (7.9)
<b>Refinement</b>		
No. of Reflections	48,862	29,814
R <sub>work</sub> / R <sub>free</sub>	0.19 / 0.23	0.18 / 0.23
No. of Atoms	4087	4102
Protein	3836	3837
Ligand/Ion	2	50
Water	249	215
R.M.S. Deviations		
Bond Lengths (Å)	0.010	0.017
Bond Angles (°)	1.24	1.8

Numbers in parenthesis represent values of highest resolution shell.

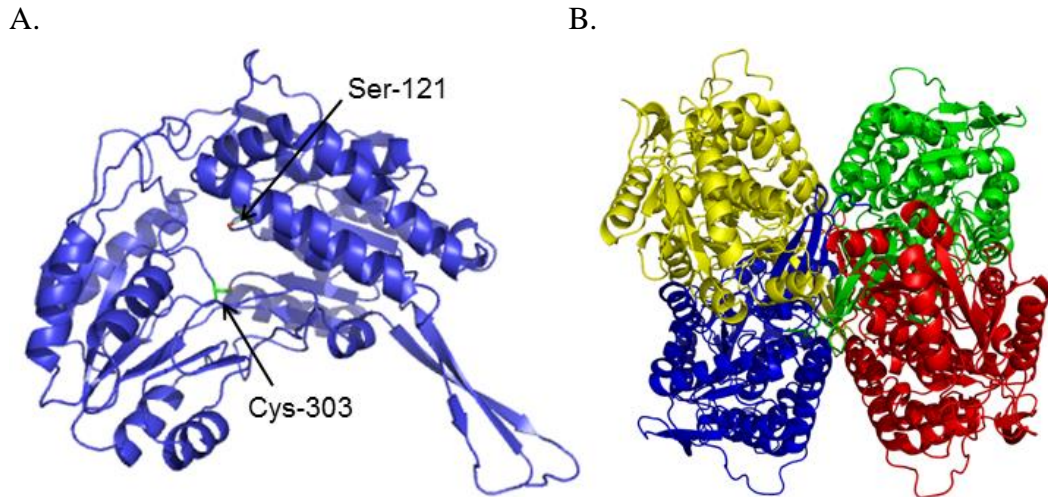


Figure 34: Structure of human ALDH1A1 (N121S) apo-enzyme. A) Ribbon representation of an ALDH1A1 monomer showing the location of cysteine 303 in the active site plus the location of ALDH1A1 N121S SNP used for the HTS. B) Ribbon representation of the structure of the homotetrameric ALDH1A1 with each monomer colored separately (PDB 4WJ9).

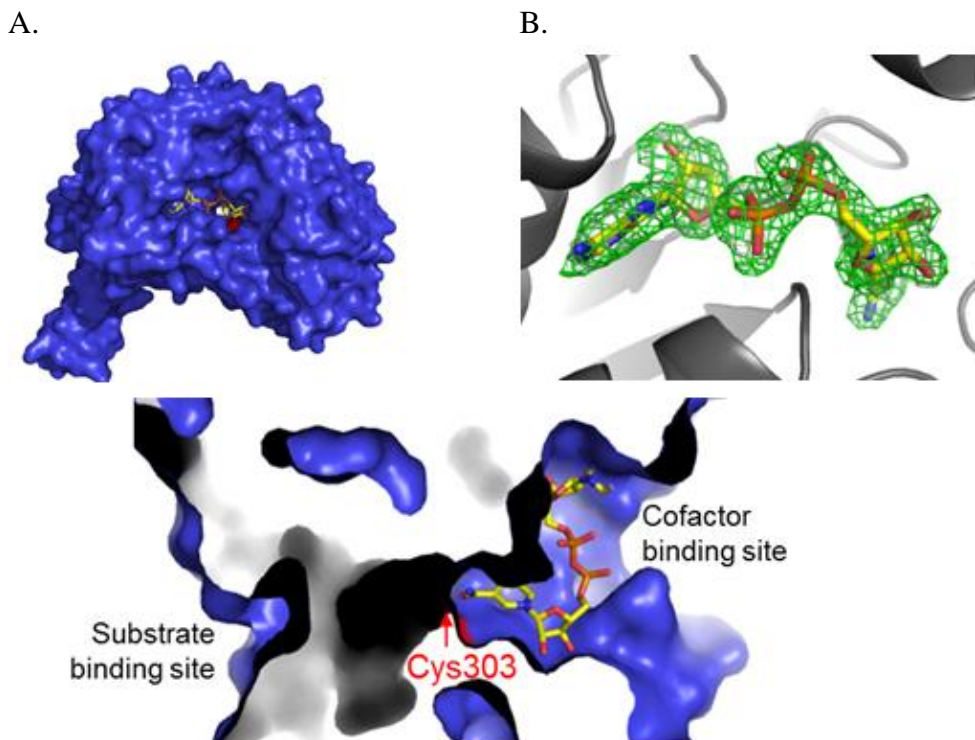


Figure 35: Structure of ALDH1A1 (WT) with reduced cofactor NADH. The location of the active site Cys-303 is shown in red. A) Location of NADH at cofactor binding site of one monomer of ALDH1A1. B) Electron density maps of NADH with the original  $F_o - F_c$  in green contoured at 2.5 standard deviations and the final  $2F_o - F_c$  map in grey contoured at 1.0 standard deviations C) Surface rendition of NADH near the active site Cys-303 (PDB 4WB9).

## 2. Esterase High-Throughput Screen

A high-throughput screen (HTS) was used to discover novel activators and inhibitors for ALDH1A1. For the HTS, we used the inherent NAD<sup>+</sup>-independent esterase activity of ALDH1A1 to identify compounds capable of altering enzyme activity, instead of the standard NAD<sup>+</sup>-dependent aldehyde oxidation assay. This approach minimized interference from compounds that absorb light at a wavelength similar to the reduced cofactor NADH. It also minimized identification of compounds that bind to the cofactor binding site. The cofactor binding site is highly conserved amongst the members of the ALDH superfamily. As seen in Figure 36, there is a high degree of similarity between the structures of ALDH1A1-NADH and ALDH2-NADH, and it is highly likely that such compounds would not be selective for ALDH1A1.

The Z-factor for the HTS comparing ALDH1A1 plus/minus inhibitor (Aldi-1) under screening conditions was 0.67 (n = 384), indicating inhibitors can be identified from single assays. As shown in Figure 37, there is a clear separation between the control reaction containing enzyme and substrate (ES Control) represented in blue, and the inhibitor control reaction containing enzyme, substrate plus an ALDH1A1 inhibitor (ESI control) represented in red. Also, the average value for ALDH1A1 with control inhibitor

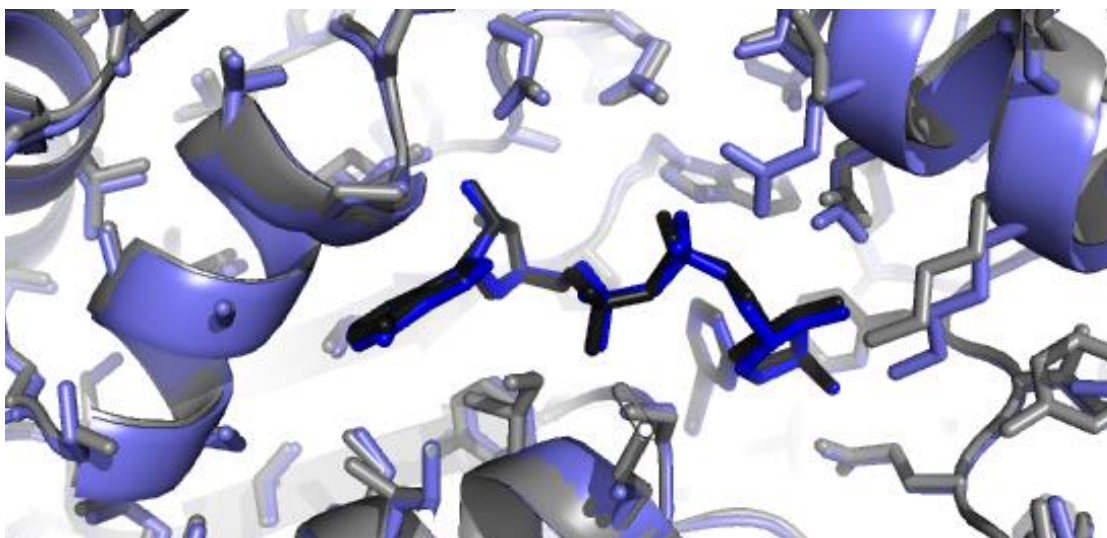


Figure 36: Comparison of the cofactor binding sites of ALDH1A1 and ALDH2. Overlap of the structure of ALDH1A1-NADH, in blue, with the structure of ALDH2-NADH, in grey.

was similar to the no enzyme, blank control (mean rate of change of 0.70 vs 0.60), indicating our inhibition control (25  $\mu\text{M}$  Aldi-1,  $\text{IC}_{50} = 2.2 \mu\text{M}^{137}$ ) strongly inhibited ALDH1A1. For the HTS, we used an ALDH1A1 protein with a known SNP at residue 121. This N121S “mutant” in the open reading frame was cloned by the Weiner group<sup>152</sup> and utilized for all their published work on ALDH1A1. The enzyme is active and behaved similarly to ALDH1A1 WT ( $K_m$  of 12  $\mu\text{M}$  vs 15  $\mu\text{M}$ , respectively, with identical  $k_{\text{cat}}/K_m$  values at  $2.7 \text{ min}^{-1} \cdot \mu\text{M}^{-1}$  for the substrate propionaldehyde). A comparison of the respective alpha-carbons in the structure of the N121S apo-enzyme and those of the wild-type ALDH1A1 structure complexed with NADH generated an RMSD of 0.2  $\text{\AA}$ , consistent with a high degree of functional and structural similarity. The side chains of Ser and Asn both form similar hydrogen bonding interactions with Tyr297. The screen used a saturating amount of the esterase substrate pNPA ( $K_m = 5 \mu\text{M}^{170}$ ). Each plate contained a control column with enzyme and substrate (ES control), and the average ( $n = 16$ ) of this intra-plate control served as the basis to determine whether a compound modified esterase activity. An activator was defined as having 2-fold or higher esterase activity compared to this control, while an inhibitor had 40% or less activity. Each plate also contained a positive control for inhibition (ESI control) using the inhibitor Aldi-1.

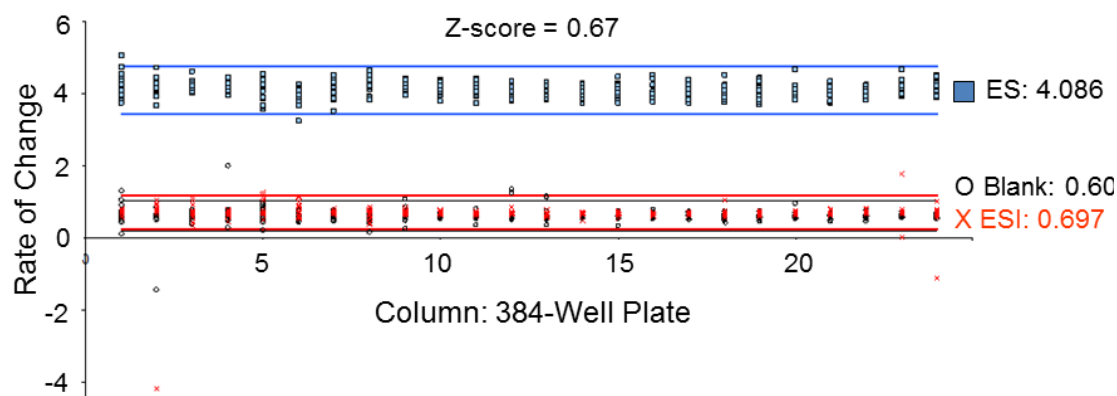


Figure 37: Z-factor determination for HTS esterase screen.

Each point represents the rate of change in absorbance at 405 nm of a reaction. The x-axis is the column (1 – 24) on the 384-well plate of the reaction. The blue data points represent the enzyme + substrate (ES) control, with an average value of 4.086; the red is enzyme + substrate + inhibitor, with an average value of 0.697; the open circles are the no enzyme control (blank). The lines represent 3x standard deviation from ES control (blue lines), ESI control (red lines), and blank (black lines). Each condition (ES, ESI, blank) performed on a separate plate with  $n = 384$ .

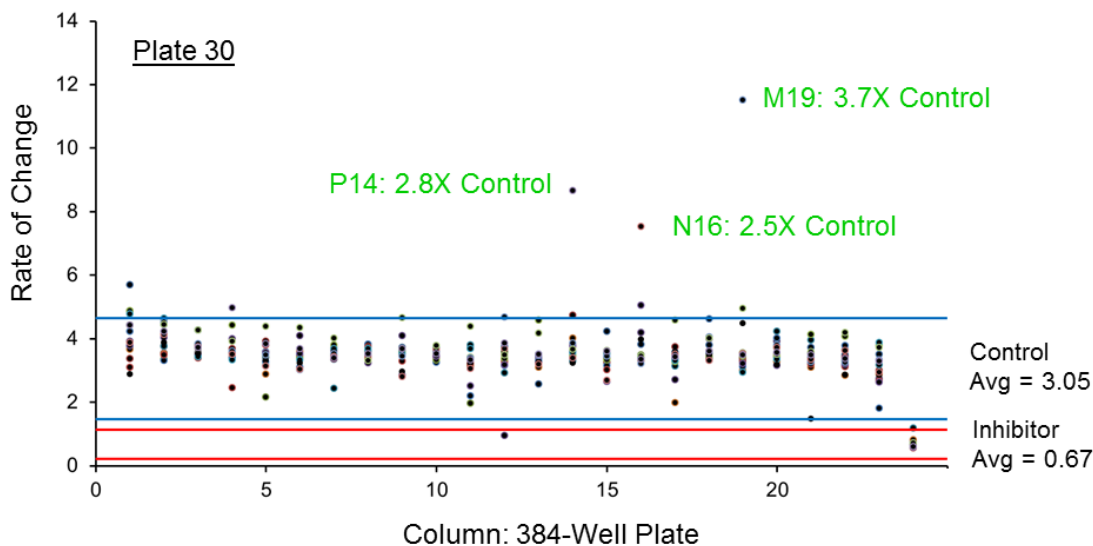


Figure 38: Representative plate from HTS.

Each point represents one well, with the x-axis the column (1-24) on the plate and the y-axis, the rate of change measured at wavelength 405 nm. Column 23 is the ES control, with an average value of 3.05 (n=16). Column 24 is the inhibition (ESI) control containing 25  $\mu$ M Aldi-1. For this plate, an activator had a value  $\geq 6.1$  while an inhibitor had a value  $\leq 1.22$ . Lines are 3x standard deviation, blue for ES and red for ESI. On this plate, we identified 3 activators (P14, N16, M19) and 1 inhibitor (D12) out of 352 compounds, with labeling based on their row and column on 384-well plate.

The initial round of the *in vitro* esterase-based screen of 64,000 compounds yielded 631 compounds that activated ALDH1A1 and 278 compounds that inhibited ALDH1A1. A sample plate from the first round of screening is shown in Figure 38, illustrating 3 activators and 1 inhibitor out of 352 compounds tested. Following rescreening of the 909 compounds identified in the first round under identical conditions, nearly 75% did not meet these same selection criteria during the second, validation assay set. After two rounds, we identified 241 activators and 15 inhibitors of ALDH1A1 esterase activity (Supplementary Tables 1 and 2).

### 3. Selectivity of Compounds for ALDH1A1 Aldehyde Oxidation

From the HTS, we identified 256 compounds that either activated or inhibited ALDH1A1 esterase activity to the levels of our selection criteria. All compounds had molecular masses between 250 and 500 Da. The 241 activators were classified into 17 groups based

on their structural similarities. Over 65% of the esterase activators fell into nine structural classes, with one class containing 78 hits. There were 22 structurally unique compounds. The HTS identified only 15 esterase inhibitors, 2 structurally similar compounds and 13 structurally unique compounds. Of the 256 hits, 67 have been analyzed further to determine their effect on aldehyde oxidation and selectivity for ALDH1A1, 57 esterase activators and 10 esterase inhibitors (Figure 39). Of the 15 esterase inhibitors identified by HTS, only 7 were commercially available without requiring customized synthesis. However, close analogs of three others were purchased and analyzed: (1) hit 3343-2924 was substituted with 2188-3302 (CM310), (2) hit C699-0615 was substituted with C699-0244 (CM306), and (3) hit K788-2754 was substituted with K938-0803 (CM307). The effects of the hits on aldehyde oxidation were tested using the standard assays performed in our laboratory to study ALDH1A1, ALDH2, and ALDH3A1. For ALDH1A1 and ALDH2, the 100  $\mu$ M propionaldehyde used is near or at saturation (ALDH1A1  $K_m$  =15  $\mu$ M and ALDH2  $K_m$  < 1  $\mu$ M). For ALDH3A1, the concentration of benzaldehyde used was set at its  $K_m$ . None of the 67 compounds activated aldehyde oxidation by ALDH1A1, ALDH2, or ALDH3A1 by more than 20%. However, of the 57 esterase activators examined at 20  $\mu$ M, 28 inhibited ALDH1A1 propionaldehyde oxidation by at least 50%. Of the 10 esterase inhibitors tested at 20  $\mu$ M, 4 compounds also inhibited ALDH1A1 propionaldehyde oxidation at least 50%, but 2 inhibitors (CM302 and CM303) also exhibited at least 50% inhibition of ALDH2 and therefore were not selective for ALDH1A1. To a limited degree, CM302 also inhibited ALDH3A1, but none of the remaining 66 hits altered ALDH3A1 benzaldehyde oxidation more than 20% from control. Based on the selectivity assays of 67 esterase hits, 30 compounds selectively inhibited ALDH1A1 compared to ALDH2 and ALDH3A1, while 2 compounds inhibited both ALDH1A1 and ALDH2 at least 50% but not ALDH3A1 (Figure 39). Therefore, the HTS for esterase activity successfully identified selective inhibitors for ALDH1A1 aldehyde oxidation.



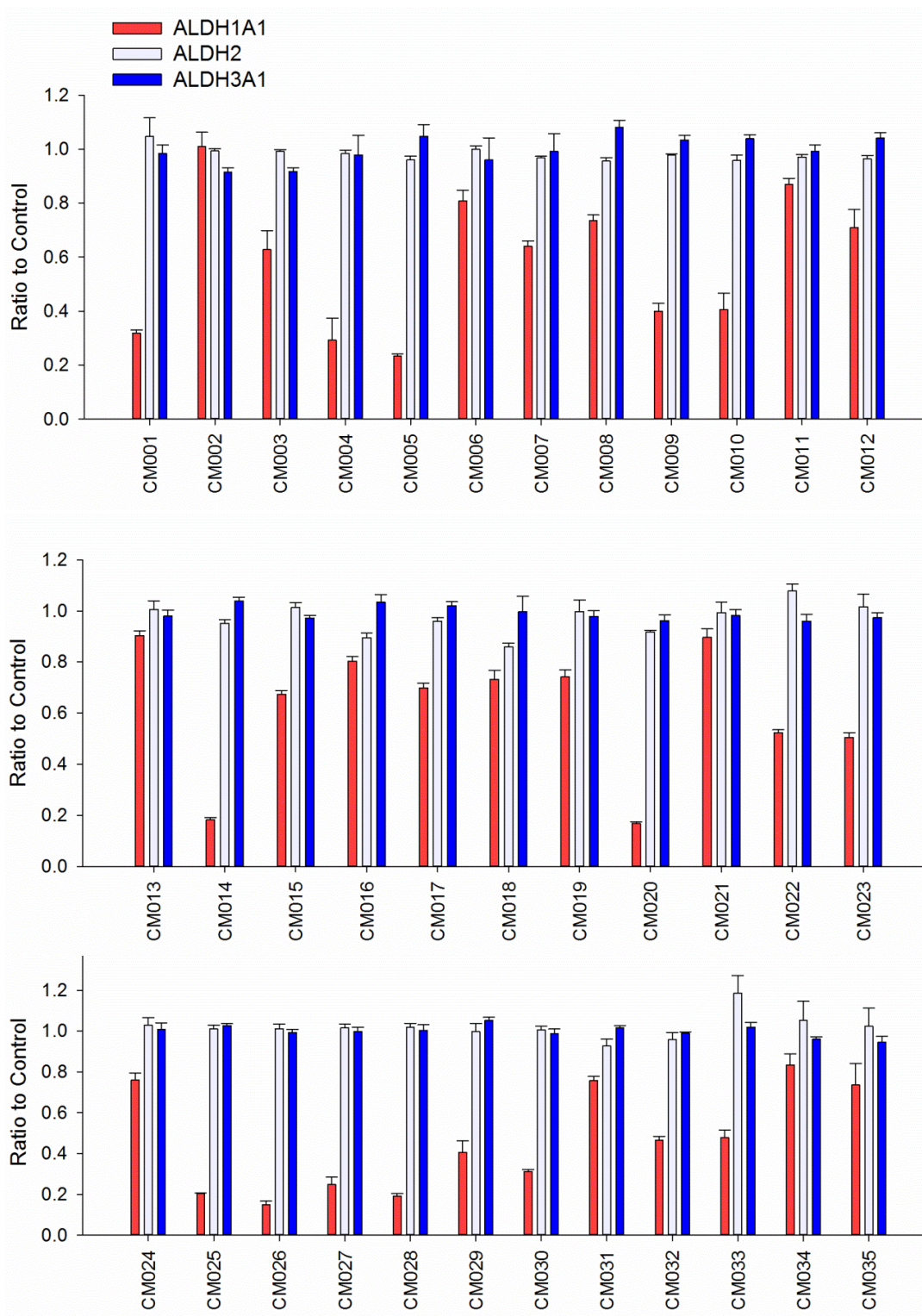


Figure 39: Effect on aldehyde oxidation of hit compounds from HTS. The activity of 67 compounds (20  $\mu\text{M}$  except [CM307] = 10  $\mu\text{M}$ ) identified via HTS was tested on three ALDH isoenzymes. Mean/SEM (n = 3)

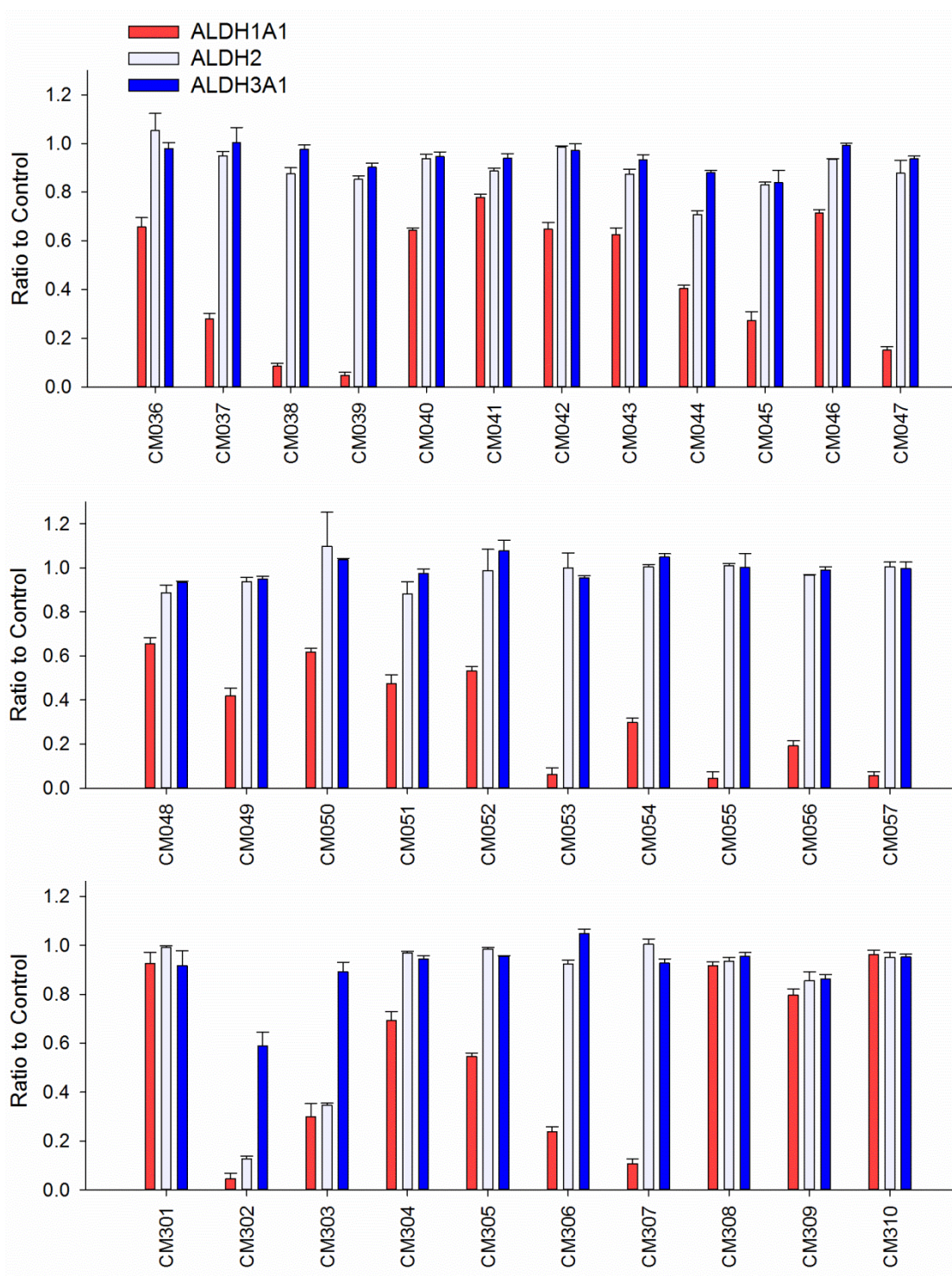


Figure 39, continued. Effect on aldehyde oxidation of hit compounds from HTS. The activity of 67 compounds (20  $\mu$ M except [CM307] = 10  $\mu$ M) identified via HTS was tested on three ALDH isoenzymes. Mean/SEM (n = 3).

IC<sub>50</sub> values were determined for compounds that inhibited propionaldehyde oxidation at least 60% at 20 μM concentration, with the most potent inhibitors and their IC<sub>50</sub> values shown in Figure 40 and Table 7. Of the 57 esterase activators, 17 were structurally similar (CM022-031, CM051-057) with all but one (CM024) inhibiting ALDH1A1 at 20 μM concentration. Based on IC<sub>50</sub> values, the most potent inhibitors selective for ALDH1A1 were CM038, CM039, and two structural analogs, CM053 and CM055, with all four hits having IC<sub>50</sub> values less than 500 nM. CM302 was a potent inhibitor of both ALDH1A1 and ALDH2, with IC<sub>50</sub> values of 1.0 ± 0.1 μM and 2.2 ± 0.3 μM, respectively. To a limited extent, CM302 also inhibited ALDH3A1, but with an IC<sub>50</sub>

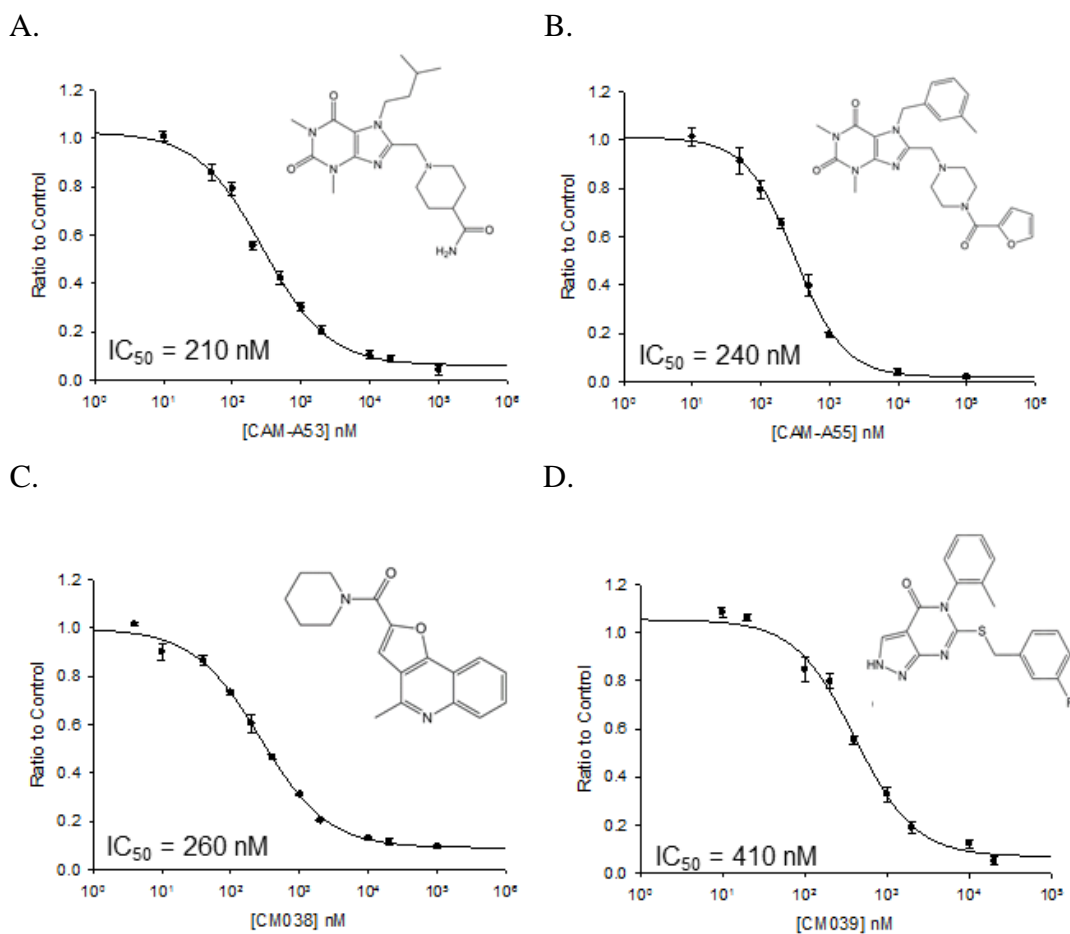


Figure 40: IC<sub>50</sub> curves and structures of the four most potent ALDH1A1 inhibitors. Representative IC<sub>50</sub> curves and structures for the three most potent ALDH1A1 inhibitors based on IC<sub>50</sub> values: A) CM053, B) CM055, C) CM038, and D) CM039. Each point represents the mean/SEM for three readings. The IC<sub>50</sub> values represent the mean of three independent assays.

value greater than 10-fold higher compared to ALDH1A1 and ALDH2. In comparison, the non-selective inhibitor Aldi-1, which was used as a control during the esterase HTS, has an  $IC_{50}$  value of 2.2  $\mu\text{M}$  for ALDH1A1<sup>137</sup>. As previously discussed, DEAB is a nonselective ALDH1/2 inhibitor used as a control for the Aldefluor Assay (Stemcell Technologies, Vancouver, Canada), a flow cytometry assay commonly used to identify stem cells based on ALDH activity. DEAB has an  $IC_{50}$  value of approximately 60 nM under these same conditions, but is also a potent inhibitor of other ALDH isoenzymes<sup>127</sup>.

Two classes of compounds identified in the HTS have been analyzed further to determine whether they could progress from hit to lead compounds in the search for selective inhibitors of ALDH1A1. CM026, in the same structural class as CM053 and CM055, and CM037 emerged from the HTS as ALDH1A1 esterase activators, but both compounds inhibited aldehyde oxidation and were selective for ALDH1A1 compared to ALDH2 and ALDH3A1. These two classes of compounds were further explored via steady-state enzyme kinetics, structural-activity relationship (SAR), X-ray crystallography, and cell culture.

#### **4. Steady-State Kinetics Optimization**

In order to perform steady-state kinetics, it was first necessary to calculate the  $K_m$  values for  $\text{NAD}^+$ , propionaldehyde, and acetaldehyde to determine the range over which to vary these substrates (Figure 41). For steady-state kinetics, their values should vary approximately 10-fold, with the  $K_m$  near the middle of the range. For  $\text{NAD}^+$ , the calculated  $K_m$  value was 50  $\mu\text{M}$ . Propionaldehyde is commonly used as a substrate for ALDH1A1 and was used for both selectivity and  $IC_{50}$  calculations. However, its calculated  $K_m$  of 15  $\mu\text{M}$  would have required varying propionaldehyde in the range of 5 – 50  $\mu\text{M}$ , a range too small to accurately detect differences in the NADH produced with our spectrophotometers. Instead, acetaldehyde ( $K_m = 180 \mu\text{M}$ ) was used, allowing for a larger range of substrate concentrations (100 – 800  $\mu\text{M}$ ) that could be monitored on a spectrophotometer.

Table 7: IC<sub>50</sub> values and structures of the top 19 compounds that inhibit ALDH1A1.

Compound IC <sub>50</sub> [μM]	Structure	Compound IC <sub>50</sub> [μM]	Structure
CM001* 1.1 ± 0.1		CM045* 2.5 ± 0.5	
CM009* 5.3 ± 0.3		CM047* 0.31 ± 0.03	
CM010* 1.3 ± 0.1		CM053 0.21 ± 0.04	
CM020 0.45 ± 0.10		CM055 0.24 ± 0.04	
CM025 2.1 ± 0.7		CM056 5.4 ± 0.8	
CM026 0.80 ± 0.06		CM057 0.92 ± 0.2	
CM028 2.0 ± 0.1		CM302 1.1 ± 0.1	
CM037 4.6 ± 0.8		CM306 3.5 ± 0.6	
CM038 0.26 ± 0.01		CM307 0.57 ± 0.09	
CM039 0.41 ± 0.01			

Each value represents mean/SEM for three independent assays, each n = 3. Values calculated using 100 μM propionaldehyde and 200 μM NAD<sup>+</sup>. An asterisk (\*) indicates less than 70% maximum inhibition.

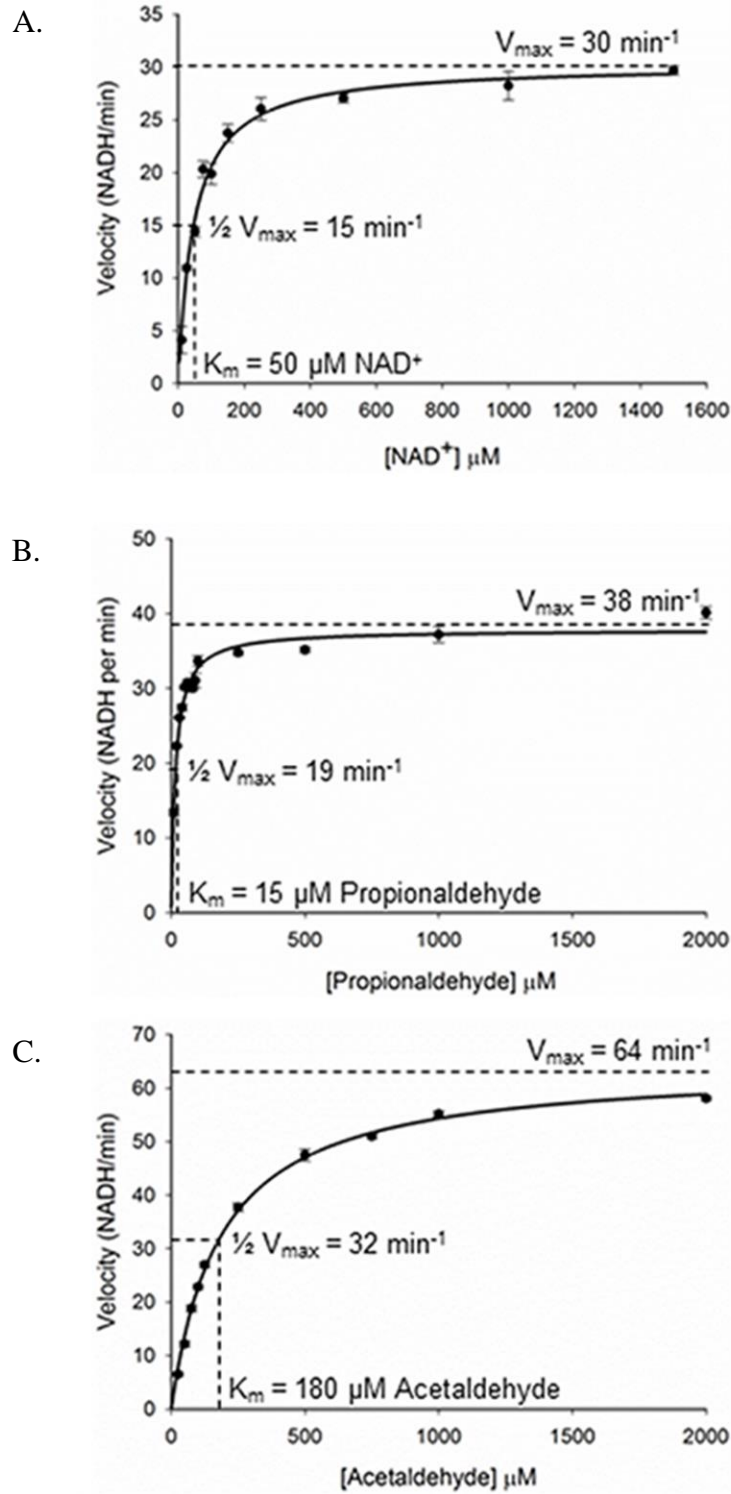


Figure 41: Calculation of  $K_m$ 's for steady-state kinetics with ALDH1A1. A)  $K_m$  of cofactor  $\text{NAD}^+$ . B)  $K_m$  of substrate propionaldehyde, and C)  $K_m$  of substrate acetaldehyde.

## 5. Characterization of the CM037 Class of Compounds

CM037 is a selective inhibitor for ALDH1A1 with a molecular weight of 431.6 Daltons (Figure 42A). At a concentration of 20  $\mu\text{M}$ , CM037 had no effect on seven other human ALDH isoenzymes (ALDH1A2, ALDH1A3, ALDH1B1, ALDH2, ALDH3A1, ALDH4A1, and ALDH5A1) as well as the carboxyl-terminal, ALDH domain of rat ALDH1L1 (Figure 42B). Higher concentrations were not tested due to the absorbance properties of CM037, which interferes at 340 nm. At 20  $\mu\text{M}$ , ALDH1A3 was inhibited approximately 25% while ALDH1A1 had 80% inhibition at this same concentration; this

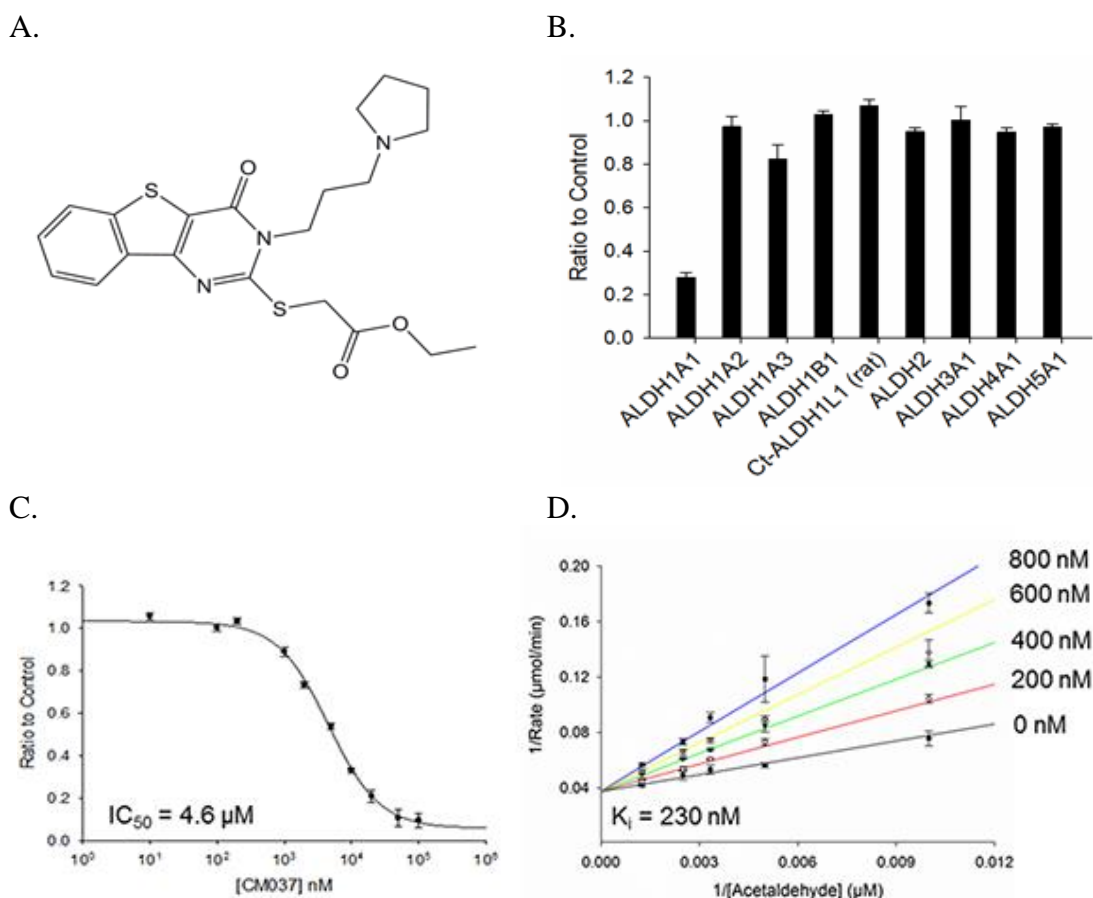


Figure 42: Characterization of CM037, a selective inhibitor of ALDH1A1.

A) The structure of CM037 with a molecular weight of 431.65 Daltons. B) Selectivity of 20  $\mu\text{M}$  of CM037 with respect to six ALDH isoenzyme. C) IC<sub>50</sub> of CM037 with ALDH1A1. D) Lineweaver-Burk representation of competitive inhibition for CM037 (0 – 800 nM) at fixed concentration of NAD<sup>+</sup> (800  $\mu\text{M}$ ) versus varied acetaldehyde (100 – 800  $\mu\text{M}$ ). The IC<sub>50</sub> curves and Lineweaver-Burk plots represent one of three experiments performed for each condition, with each point the mean/SEM of three data points at each concentration.

concentration had little to no effect on the other isoenzymes. CM037 had good potency towards ALDH1A1 with an  $IC_{50} = 4.6 \pm 0.8 \mu\text{M}$  (Figure 42C). CM037 has a competitive mode of inhibition with respect to varied acetaldehyde, with a  $K_i$  of  $0.23 \pm 0.06 \mu\text{M}$  (Figure 42D).

Although it had some structural similarity to CM039, no analogs of CM037 were identified in the HTS. To better understand the role of particular chemical groups, the selectivity and potency of ten compounds structurally similar to CM037 were tested. Of particular interest was the role of CM037's ester group. As an ester, it is possible CM037 is a substrate for ALDH1A1, and

in a manner similar to what was seen with DEAB, the ester substrate is turned over at such a slow rate that it inhibits aldehyde oxidation. SAR was performed using CM037 analogs to determine the effect of structural changes to both selectivity and potency. CM037a had an identical R1 group but the ester group of R2 was replaced with a primary amide. The  $IC_{50}$  value for the amide is 5 times higher than the ester ( $23 \mu\text{M}$  vs  $4.6 \mu\text{M}$ ) (Figure 44), indicating that an amide could not replace the ester functional group. Nine of the analogs contained the identical fused three-ring structure as CM037, with various R1 and R2 groups (Figure 45), while the tenth analog was similar to CM039. None of the ten compounds had any effect on ALDH3A1. However, there was great variability in the effect these compounds had on the three ALDH1A subfamily members (Figure 45). Four analogs had identical R2 groups to CM037, maintaining the ester group (CM037b, CM037c, CM037g, and CM037j). At a compound concentration of  $20 \mu\text{M}$ , only one compound (CM037g) inhibited ALDH1A1 more than 50%. CM037g replaces the propyl pyrrolidine with an ethanol. This change had little effect on potency towards ALDH1A1, having an  $IC_{50}$  only slightly lower than CM037 ( $3.3 \mu\text{M}$  vs.  $4.6 \mu\text{M}$ , Figure 44). However, this R1 change resulted in ALDH1A3 inhibition increasing from 20% with CM037 to 70% with CM037g, indicating that the R1 group is involved in isoenzyme selectivity with respect to ALDH1A3. For the other three esters, the pyrrolidine group

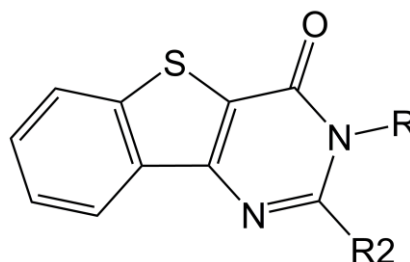


Figure 43: Basic structure of CM037 Analogs.



three carbons removed from the core structure was replaced with a phenyl group closer to the core structure. At 20  $\mu\text{M}$  concentration, none of these three compounds inhibited ALDH1A1 more than 50%, indicating a longer linker may be needed to inhibit ALDH1A1. However, the opposite was seen with the other members of the ALDH1A subfamily as both CM037b and CM037j had an increase in inhibition of ALDH1A2 and ALDH1A3. Only one non-ester analog (CM037e), with a carboxylic acid group at the R1 position, inhibited ALDH1A1 more than 50% at the tested concentration.

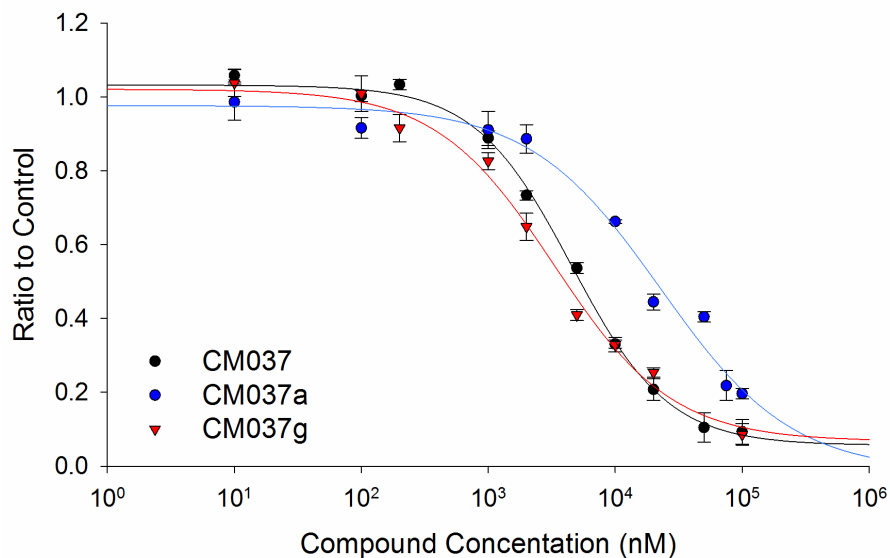


Figure 44: IC<sub>50</sub> Curves for CM037 and two of its analogs.

The parent compound CM037, an ester, is shown in black ( $\text{IC}_{50} = 4.6 \mu\text{M}$ ). An ester analog CM037g, shown in red, is slightly more potent than the parent compound ( $\text{IC}_{50} = 3.3 \mu\text{M}$ ). The amide CM037a, shown in blue, is less potent ( $\text{IC}_{50} = 23 \mu\text{M}$ ). Each point represents the mean/SEM of three readings, with the  $\text{IC}_{50}$  values the mean/SEM of three independent experiments.

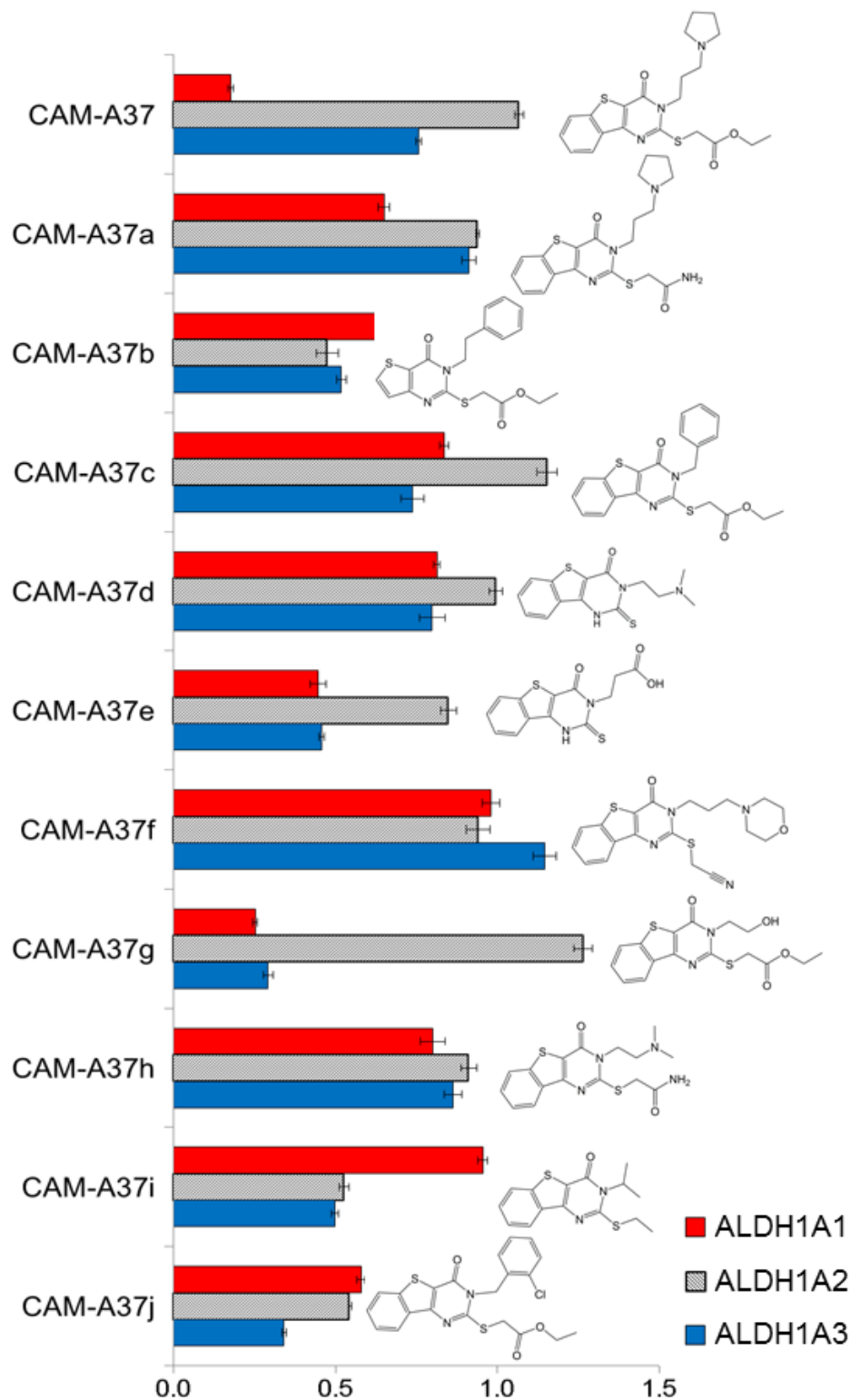


Figure 45: CM037 and its analogs.

The effect on aldehyde oxidation activity of ten CM037 analogs on three ALDH1A isoenzymes. The reactions used 20  $\mu\text{M}$  compound and each bar represents mean/SEM (n = 3).

## 6. Characterization of the CM026 Class of Compounds

CM026 is a selective inhibitor for ALDH1A1 with a molecular weight of 442.5 Daltons (Figure 46).

At a concentration of 20  $\mu\text{M}$ , CM026 had no effect on seven other human ALDH isoenzymes (ALDH1A2, ALDH1A3, ALDH1B1, ALDH2, ALDH3A1, ALDH4A1, and ALDH5A1) as well as the carboxyl-terminal ALDH domain of rat ALDH1L1 (Figure 47A). At a concentration of 100  $\mu\text{M}$  compound, CM026 modestly increased

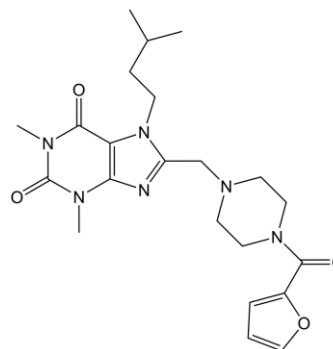


Figure 46: Structure of CM026.

aldehyde oxidation catalyzed by ALDH1A2, ALDH1A3, and ALDH1B1, but 20  $\mu\text{M}$  CM026 had values similar to control. For an initial hit compound, CM026 had good potency towards ALDH1A1 with an  $\text{IC}_{50} = 0.80 \pm 0.06 \mu\text{M}$  (Figure 47B). Complete inhibition of ALDH1A1 was not observed, with concentrations above 20  $\mu\text{M}$  leveling off between 80 and 90% inhibition. CM026 has a noncompetitive partial mode of inhibition with respect to varied acetaldehyde, with a  $K_i$  of  $0.80 \pm 0.16 \mu\text{M}$  and  $\beta = 0.15 \pm 0.03$  (Figure 48A), indicating that maximum inhibition was  $0.15V_{\text{max}}$ . With respect to varied  $\text{NAD}^+$ , CM026 had an uncompetitive partial mode of inhibition, with a  $K_i$  of  $0.72 \pm 0.03 \mu\text{M}$  and  $\beta = 0.10 \pm 0.03$  (Figure 48B).

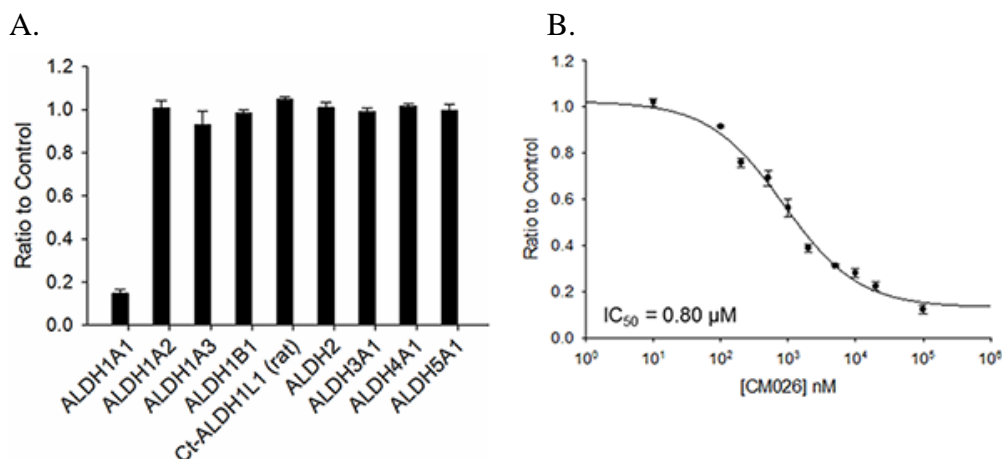


Figure 47: Selectivity and potency of CM026.

A) Selectivity of 20  $\mu\text{M}$  CM026 with respect to nine ALDH isoenzymes. B).  $\text{IC}_{50}$  curve of CM026 with ALDH1A1. The  $\text{IC}_{50}$  curve represent one of three experiments performed, with each point the mean/SEM of three data points at each concentration.

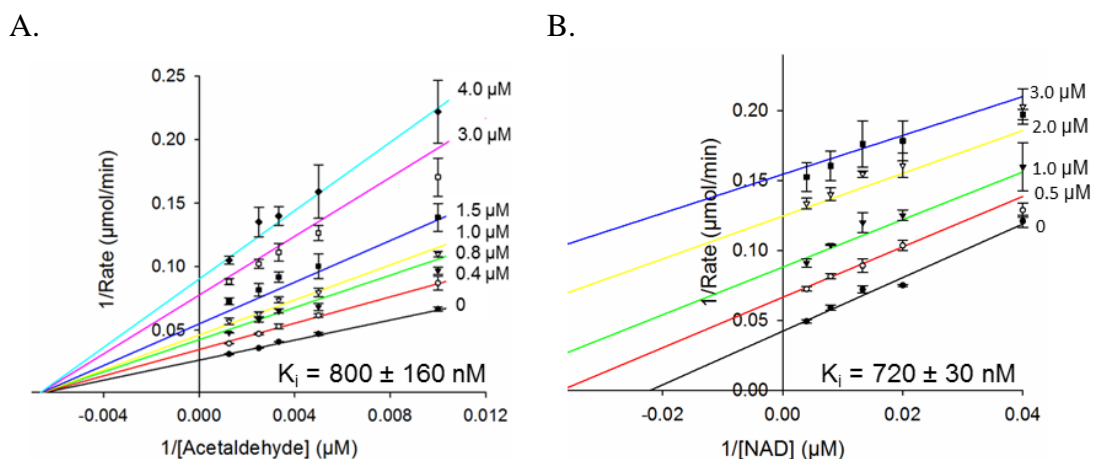


Figure 48: Mode of Inhibition for CM026, a selective inhibitor of ALDH1A1. A) Lineweaver-Burk representation of noncompetitive inhibition for CM026 (0 - 3  $\mu\text{M}$ ) versus varied acetaldehyde (100 - 800  $\mu\text{M}$ ) at fixed concentration of  $\text{NAD}^+$  (800  $\mu\text{M}$ ). B) Lineweaver-Burk representation of uncompetitive inhibition for CM026 (0 - 3  $\mu\text{M}$ ) versus varied  $\text{NAD}^+$  (25 - 200  $\mu\text{M}$ ) at fixed concentration of propionaldehyde (200  $\mu\text{M}$ ). Lineweaver-Burk plots represent one of three experiments performed, with each point the mean/SEM of three data points at each concentration.

The HTS identified 77 compounds that were structurally similar to CM026, with the basic xanthine structure shown in Figure 49. These compounds share this core structure with two commonly used compounds, theophylline and caffeine. However, neither theophylline nor caffeine affected ALDH1A1 aldehyde oxidation activity at concentrations up to 250  $\mu\text{M}$ , indicating that larger substituents at the R1 and/or R2 positions are necessary for ALDH1A1 inhibition (Figure 50). Seventeen members of this class of compounds have been tested further, enabling us to examine their structural activity relationship (SAR). Unlike traditional SAR, because these compounds were identified in the HTS, all 17 hits were known to bind to ALDH1A1 and affect the enzyme's esterase activity. All inhibited ALDH1A1 to some degree while having little to no effect on ALDH1A2, ALDH1A3, ALDH1B1, ALDH2, and ALDH3A1 (Table 8). At 100  $\mu\text{M}$ , some compounds activated

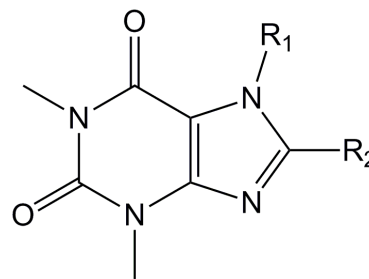


Figure 49: General structure of CM026 analogs.

the aldehyde oxidation of some of the other three ALDH1 isoenzymes tested, similar to the activation seen with CM026, but at 20  $\mu\text{M}$  any activation was less than 10%. CM053 was the most potent analog examined, with an  $\text{IC}_{50} = 210 \pm 40 \text{ nM}$  and a  $K_i = 96 \pm 14 \text{ nM}$ , with noncompetitive tight inhibition compared to varied substrate acetaldehyde. Unlike CM026, complete inhibition was observed with CM053. CM028 shares the same R2 group as CM053 but the isopentyl group at R1 has been replaced with a phenylpropyl group. CM028 is less potent, with an  $\text{IC}_{50} = 2.0 \pm 0.1 \mu\text{M}$  and exhibits a competitive mode of inhibition with respect to acetaldehyde with a  $K_i = 240 \pm 40 \text{ nM}$ .

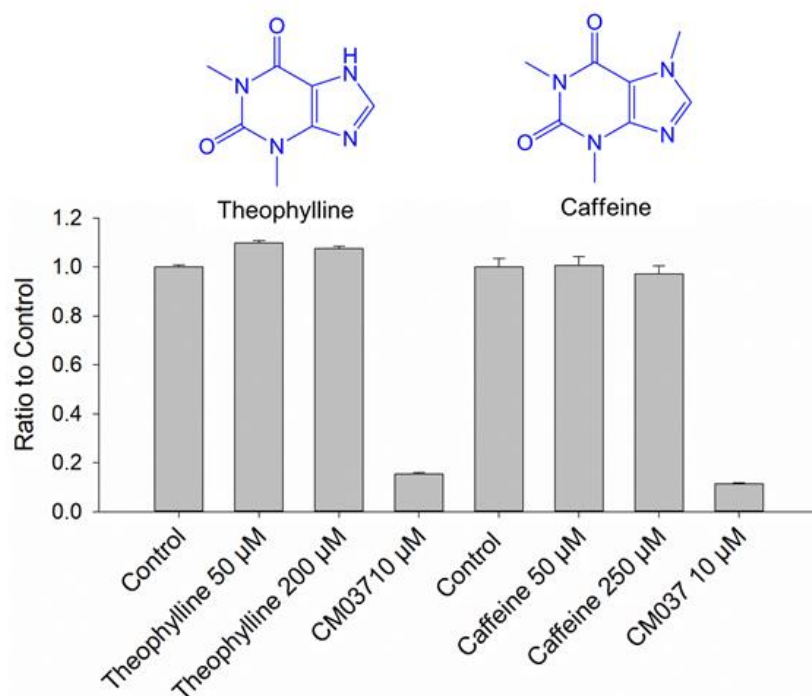
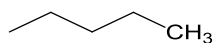
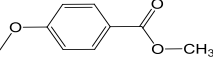
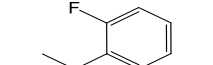
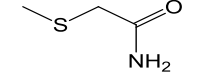
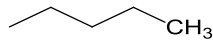
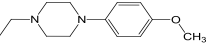
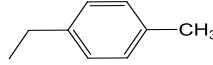
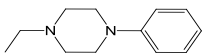
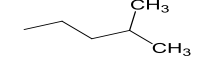
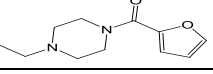
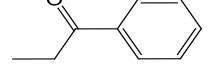
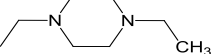
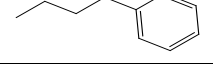

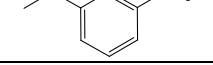
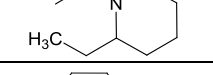
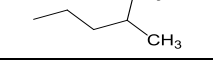
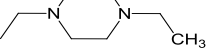
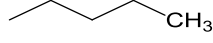
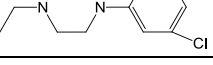
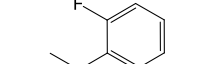
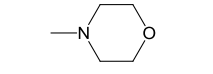
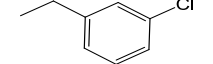
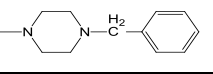
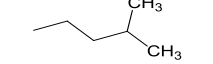
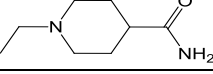
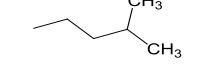
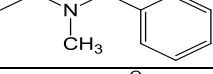
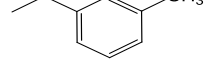
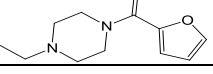
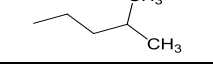
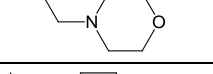
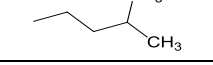
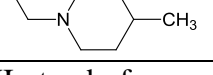


Figure 50: Common structural analogs of CM026. The structures and effects of theophylline and caffeine on ALDH1A1. Each value is mean/SEM, with  $n = 3$ .

Table 8: SAR for CM026 Analogs.

	R1	R2	IC <sub>50</sub> in $\mu\text{M}$ (SE)					
			ALDH 1A1	ALDH 1A2	ALDH 1A3	ALDH 2	ALDH 1B1	ALDH 3A1
CM022			5.2* (0.8)	>100	NI(A)	NI	NI	NI
CM023			14* (2)	NI	NI(A)	NI	NI	>100
CM024			>20	NI	>100	NI	>100	NI
CM025			2.1 (0.7)	NI	NI(A)	NI	>100	NI(A)
CM026			0.80 (0.06)	NI(A)	NI(A)	NI	NI(A)	NI(A)
CM027			6.1 (1.1)	NI	NI	NI	NI	NI
CM028			2.0 (0.1)	NI	NI	>100	>100	NI
CM029			8.4 (1.0)	NI	NI(A)	NI	NI	NI
CM030			>20	NI	NI	NI	NI	NI
CM031			>20	>100	>20	>100	>100	NI(A)
CM051			>20	NI(A)	NI(A)	NI	NI	NI
CM052			>20	>100	NI(A)	NI	NI(A)	NI(A)
CM053			0.21 (0.04)	NI	NI(A)	NI	NI	NI(A)
CM054			3.4* (0.7)	>100	NI(A)	>100	>100	NI
CM055			0.24 (0.04)	NI(A)	NI(A)	NI	NI	NI
CM056			5.8 (1.2)	NI	NI	>100	>100	NI
CM057			0.92 (0.2)	NI	>100	>100	>100	NI

At 100  $\mu\text{M}$  compound, NI stands for no inhibition and NI(A) indicates no inhibition but activation. \* indicates < 70% maximum inhibition.

## 7. X-ray Crystallography of ALDH1A1 with CM026, CM053, and CM037

To determine the mechanism that underlies the ability of these compounds to selectively inhibit ALDH1A1, X-ray crystallography was used to determine the structure of the enzyme-compound complexes. For CM026 and CM037, the naturally occurring N121S polymorphic variant of ALDH1A1 was used<sup>167</sup>, while ALDH1A1 WT was used for the structure with bound CM053. A comparison of the respective alpha carbons in the structure of the N121S·CM026 to those in the WT·CM053 generated an RMSD of 0.12Å, indicating a high degree of similarity between WT and the N121S mutant as expected since the two have very similar kinetic behavior<sup>167</sup>. For CM026 analogs, apo-ALDH1A1 crystals were soaked overnight in a crystallization solution containing 500 µM compound and 2% DMSO. Structures of the ALDH1A1-CM026 complex and the ALDH1A1-CM053 complex were solved to resolutions of 1.80 Å and 1.95 Å, respectively (Table 9).

Table 9: Data collection and refinement statistics.

<b>Data Collection</b>	<b>ALDH1A1- CM026 PDB 4WP7</b>	<b>ALDH1A1- CM053 PDB 4WPN</b>	<b>ALDH1A1- CM037 PDB 4X4L</b>
Space Group	P422	P422	P422
Cell Dimensions			
a, b, c (Å)	109, 109, 83	109, 109, 83	109, 109, 83
A, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	50 – 1.80	50 – 1.95	50 – 1.85
R <sub>merge</sub>	0.082(0.59)	0.11(0.66)	0.058 (0.70)
I/σ <sub>i</sub>	22.7 (4.7)	18.3 (3.7)	27.4 (3.3)
Completeness (%)	99 (100)	99 (100)	99 (100)
Redundancy	11.7 (11.7)	9.3 (6.9)	8.5 (8.8)
<b>Refinement</b>			
No. of Reflections	44544	35048	40517
R <sub>work</sub> / R <sub>free</sub>	0.19 / 0.22	0.19 / 0.24	0.20 / 0.23
No. of Atoms	4109	4080	4066
Protein	3833	3858	3806
Ligand/Ion	35	33	80
Water	241	189	180
R.M.S. Deviations			
Bond Lengths (Å)	0.008	0.009	0.010
Bond Angles (°)	1.28	1.30	0.135

Numbers in parenthesis represent values of highest resolution shell.

CM026 binds near the solvent exposed exit of the substrate-binding site (Figure 51, PDB Code 4WP7). The xanthine ring lies parallel to and approximately 3.6 Å from Tyr297, creating a hydrophobic pi-stacking interaction. Four residues form hydrogen bonds with CM026; the xanthine ring interacts with His293, Cys302, and Gly458, while Trp178 interacts with the ketone group on R2. The isopentyl group of R1 projects towards Cys303 and fills the hydrophobic space bounded by Phe171 and Phe466. CM053 differs from CM026 only in its R2 group, which can form hydrogen bonds with two residues, Trp178 and Val460 (Figure 52, PDB Code 4WPN). For CM037, apo-ALDH1A1 crystals were soaked for 5 hours in a solution containing 500 μM compound and 1 mM NAD<sup>+</sup> with 1% DMSO. CM037 binds at a similar location to the CM026 compounds, but its

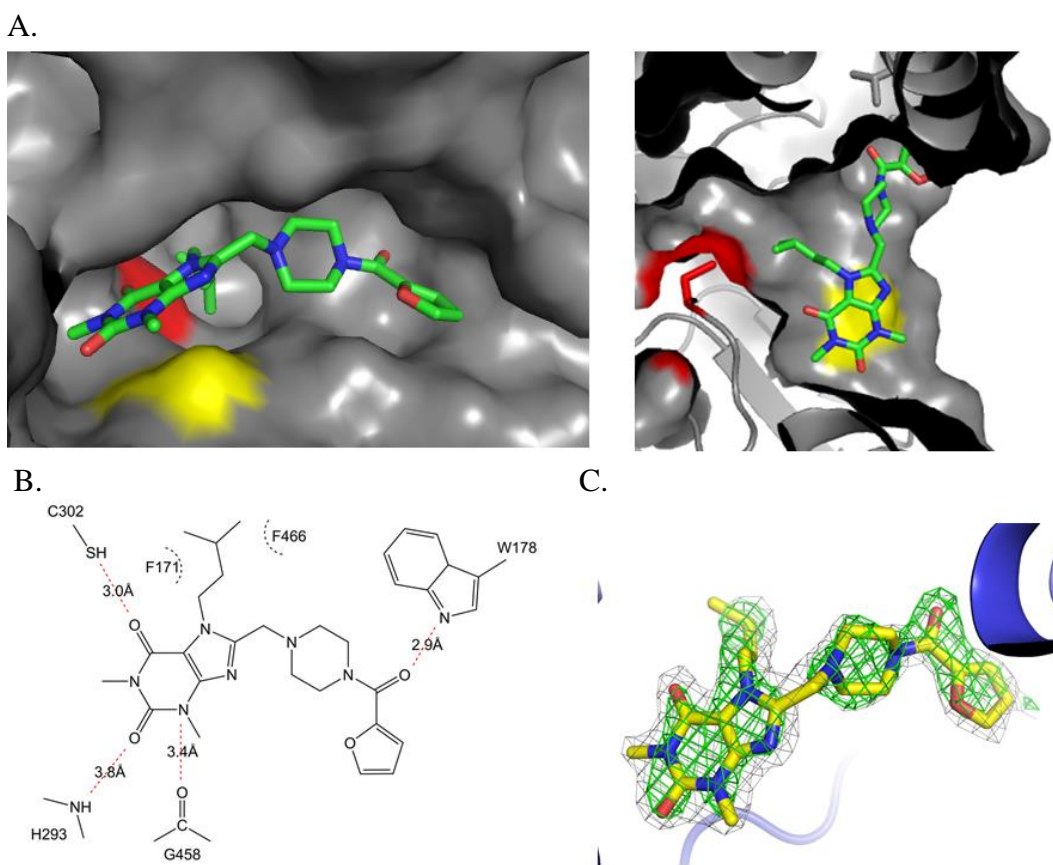


Figure 51: Structure of ALDH1A1 with CM026.

A) CM026 binds in the active site near cysteine 303, shown in red. The location of Gly458 is shown in yellow. B) Two-dimensional representation of the key hydrogen bonds, illustrated with red dashed lines, and hydrophobic interactions, illustrated with black arcs, between ALDH1A1 and CM026. C) The electron density maps of CM026, with the original  $F_o - F_c$  map in green contoured at 2.5 standard deviations and the final  $2F_o - F_c$  map in grey contoured at 1.0 standard deviations (PDB Code 4WP7).



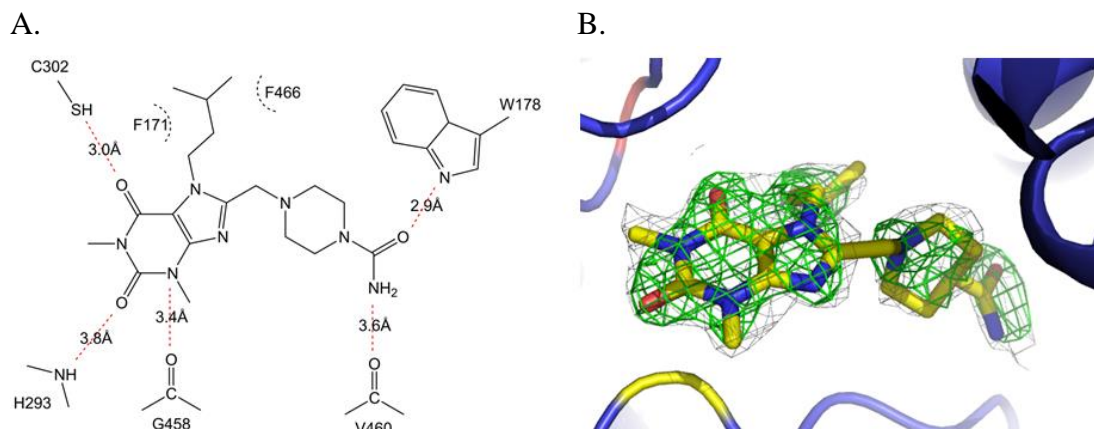


Figure 52: Structure of human ALDH1A1 with CM053.

A) Two dimensional representation of the key hydrogen bonds, illustrated with red dashed lines, and hydrophobic interactions, illustrated with black arcs, between ALDH1A1 and CM053. B) The electron density maps of CM053 with the original  $F_o - F_c$  map in green contoured at 2.5 standard deviations and the final  $2F_o - F_c$  map in grey contoured at 1.0 standard deviations (PDB Code 4WPN).

long axis is oriented almost orthogonal to that of CM026 and CM053 (Figure 53A). Most of the tricyclic ring of CM037 is in a hydrophobic pocket formed by Phe171, Val460, and Phe466 with a potential hydrogen bond between the ring system's carbonyl oxygen atom and the side chain of Cys302 (Figure 53B). The biggest structural adaptation to CM037 binding is the movement of Trp178 away from the substrate-binding site to accommodate the benzyl ring of CM037 (Figure 53C). This conformational movement appears to be dynamic and impacts the observed electron density for both the benzyl group of CM037 and of Trp178. Trp178 is well ordered in all other determined structures of human ALDH1A1, including our CM026 and CM053 structures, but has weak density for the benzyl moiety of the indole ring in this complex (Figure 53D). We would suggest that optimization of CM037 could be achieved by altering the thiophene and benzyl ring systems to alleviate these steric conflicts.

To better understand the selectivity of these compounds for ALDH1A1, we compared the structure of human ALDH1A1 against human ALDH2 (PDB code 1CW3), human ALDH3A1 (PDB code 3SZA), and human ALDH4A1 (PDB code 3V9G) and identified a critical glycine (Gly458) that is present near the xanthine ring binding site in ALDH1A1. This glycine is replaced by larger amino acid side chains in the other three structures

examined, as well as in sheep ALDH1A1. In rat ALDH1A2 (PDB code 1BI9), which shares 97% sequence identity to human ALDH1A2, this location is part of a small disordered loop not observed in the crystal structure<sup>171</sup>. Using sequence alignments of the human genes, Gly458 in ALDH1A1 is replaced by an asparagine in ALDH1A2, ALDH1A3, and ALDH1B1, an aspartate in ALDH2, and an isoleucine in ALDH3A1 (Figure 54A). As shown in Figure 54B, these side chains would interfere with the position of the xanthine ring, effectively eliminating the ability of these analogs to bind to any human isoenzyme but ALDH1A1.

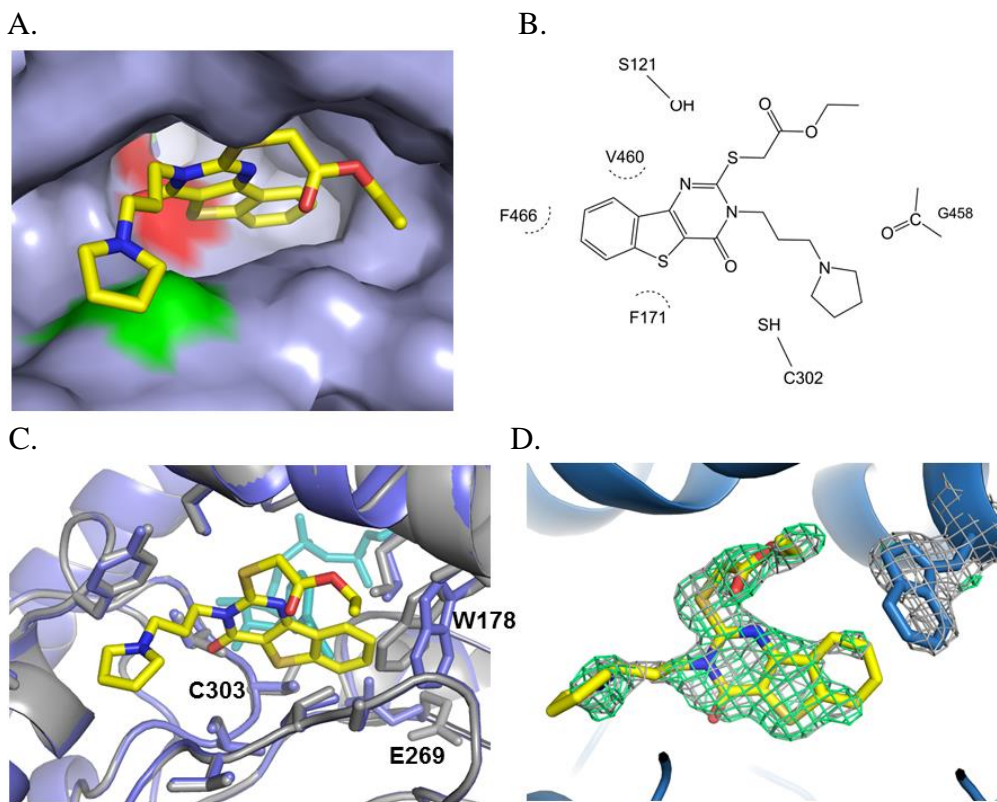


Figure 53: Structure of ALDH1A1 N121S with CM037.

(A) CM037 binds in the active site near cysteine 303, shown in red. The location of Gly458 is shown in green. (B) Two-dimensional representation of the hydrophobic interactions, illustrated with black arcs, between ALDH1A1 and CM037. (C) Binding of CM037 induces structural changes in ALDH1A1 (in blue) compared to apo-ALDH1A1 (in gray), particularly at W178. NADH binding (in cyan), induces conformational changes at the cofactor binding site, as seen here with E269. (D) The electron density maps of CM037, with the original  $F_o - F_c$  map in green contoured at 2 standard deviations and the final  $2F_o - F_c$  map in grey contoured at 1.0 standard deviations. Figures A, C and D generated in Pymol and Figure B in ChemDraw.

A.

ALDH2	TYGLAAAVFTKDLDKANYLSQAL--Q--AGTVWVNCYDVFGAQS--PFGGY	485
ALDH1A1	FYGLSAGVFTKDIDKAITISSAL--Q--AGTVWVNCYGVVSAQC--PFGGF	469
ALDH1A2	DFGLVAAVFTNDINKALTVSSAM--Q--AGTVWVNCYNALNAQS--PFGGF	486
ALDH1A3	DYGLTAAVFTKNLDKALKLASAL--E--SGTVWVNCYNALYAQA--PFGGF	480
ALDH1B1	RYGLAAAVFTRDLDKAMYFTQAL--Q--AGTVWVNTYNIVTCHT--PFGGF	485
ALDH1L1	EFGLASGVFTRDINKALYVSDKL--Q--AGTVFVNTYNKTDVAA--PFGGF	885
ALDH1L2	EYGLASGVFTRDINKAMYVSEKL--E--AGTVFINTYNKTDVAA--PFGGV	896
ALDH3A1	EKPLALYMFSSNDKVIKKMIAET--S--SGGVAANDVIVHITLHs1PFGGV	405
ALDH3A2	EKPLALYVFSHNHKLIKRMIDET--S--SGGVTGNDVIMHFTLNs1PFGGV	402
ALDH3B1	EKPLALYAFSNSSQVVKRVLtQT--S--SGGFCGNDGFMHMTLAs1PFGGV	405
ALDH3B2	EKPLALYAFSNSSQVNVQMLERT--S--SGSFGGNEGFTYISILLsvPFGGV	324
ALDH4A1	SYGLTGAVFSQDKDVVQEA TKVL--RnaAGNFYINDKSTGSIVGqgPFGGA	523
ALDH5A1	DVGLAGYFYSQDPAQIWRVAEQ--E--VGMVGVNEGLISSVEC--PFGGV	520
ALDH6A1	PYNGTAIFTTNGATARKYAHLV--D--VGQVGVNVP1PVPLPMf-SFTGS	487
ALDH7A1	KQGLSSSIFTKDLGRIFRWLGPKgSD--CGIVNVNIPtSGAEIGg-AFGGE	499
ALDH8A1	KYGLAATVWSSNVGRVHRVAKKL--Q--SGLVWtNCWLIRELNL--PFGGM	460
ALDH9A1	TFGLAAGVFTRDIQRAHRVVAEL--Q--AGTCFINNYNVSPVEL--PFGGY	484
ALDH16A1	PRGGSASVWSERLQALELGYGL--Q--VGTWVINAHGRLRDPsV--PTGGC	467
ALDH18A1	PVGLELLTtTKWLLRGKDHVVsdfS Eh--GSLKYLHENLPIQRntN----	795

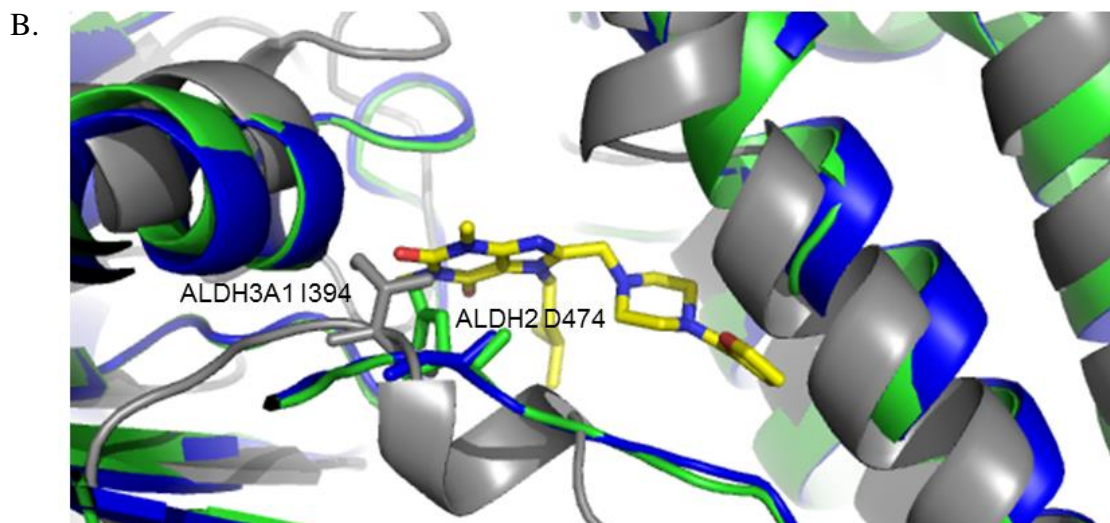


Figure 54: Structural basis of selectivity of CM026 for ALDH1A1.

A) Multiple sequence alignment in the region of ALDH1A1 Gly-458 to the mature form of ALDH2. B) Structure of ALDH1A1 (blue) with bound CM026 compared to ALDH2 (green) and ALDH3A1 (grey) indicating that a bulky amino acid such as the Asp of ALDH2 and Ile of ALDH3A1 would clash with CM026 and prevent the compound from inhibiting the enzyme. Sequence alignment was performed using NCBI delta-BLAST, while structural alignment was performed using least square fit (LSQ) in Coot.

## 8. Characterization of ALDH1A1 G458N Mutant

To confirm whether Gly458 in ALDH1A1 directly impacts the selectivity of CM026 analogs for ALDH1A1, we mutated the glycine at this position to asparagine, as found in ALDH1A2 and ALDH1A3. We determined the kinetic parameters for acetaldehyde oxidation for both the wild-type and G458N enzymes (Table 10). This mutation did not dramatically affect the enzyme's catalytic efficiency for aldehyde oxidation. However, when Gly458 is mutated to asparagine, CM026 no longer inhibits the enzyme at concentrations up to 100  $\mu\text{M}$ , and none of the CM026 analogs inhibited the mutant more than 25% (Figure 55). CM037 also does not inhibit G458N. However, DEAB and CM302, a non-selective inhibitor identified in the HTS, both inhibit G458N, exhibiting  $\text{IC}_{50}$  values of  $0.52 \pm 0.10 \mu\text{M}$  and  $3.1 \pm 0.3 \mu\text{M}$  respectively, compared to  $0.057 \pm 0.005 \mu\text{M}$  and  $1.0 \pm 0.1 \mu\text{M}$ , respectively for ALDH1A1 WT (Table 10). These data support the hypothesis that the substrate-binding site and in particular Gly458 determine the selectivity of the CM026 class of compounds for ALDH1A1, and that the bulkier side chains at this position in ALDH1A2, ALDH1A3, ALDH1B1, ALDH2, and ALDH3A1 occludes their binding to other ALDH isoenzymes.

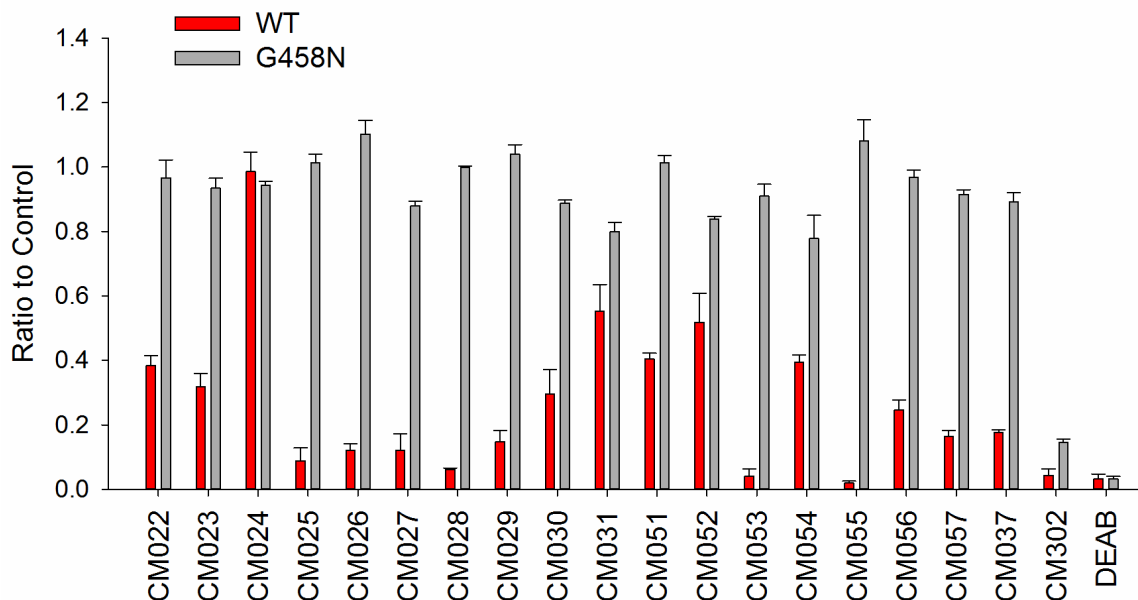


Figure 55: Selectivity of compounds for WT vs G458N mutant.

For CM026 and its analogs, 100  $\mu\text{M}$  of compound was used. For CM037 and CM302, 20  $\mu\text{M}$  of compound was used. Each value is mean/SEM, with  $n = 3$ .

Table 10: Kinetic parameters of ALDH1A1 WT and mutant G458N.

$K_M$	Acetaldehyde	$k_{cat} / K_M$	$K_I^{CM026}$	$K_I^{CM037}$	$K_I^{CM302}$	$K_I^{DEAB}$
	( $\mu M$ )	( $\text{min}^{-1} \cdot \mu M^{-1}$ )	( $\mu M$ )	( $\mu M$ )	( $\mu M$ )	( $\mu M$ )
WT	177 $\pm$ 19	0.18 $\pm$ 0.02	0.80 $\pm$ 0.16	4.6 $\pm$ 0.8	1.0 $\pm$ 0.1	0.057 $\pm$ 0.005
G458N	86 $\pm$ 2	0.21 $\pm$ 0.02	NI	NI	3.1 $\pm$ 0.3	0.52 $\pm$ 0.10

## B. Discussion

Aldehyde dehydrogenases contribute to a variety of biological processes and disease states. ALDH1A1, in particular, has been linked to such diverse diseases as cancer, Parkinson's disease, obesity, and cataracts. Therefore, selective inhibitors of ALDH1A1 would be of tremendous value to understanding the roles of this enzyme in both normal and disease processes. However, great structural and functional similarity exists, especially within the ALDH1/2 family, with five members sharing approximately 70% protein sequence identity plus significant overlap in substrate utilization. To date, there are no ALDH1A1-selective inhibitors commercially available. Although comparisons of available ALDH structures indicate a high degree of overlap, distinct surface topographies exist that may enable selective inhibitors to be developed. Comparison of the structures of human ALDH1A1, ALDH2, and ALDH3A1 indicate they exhibit a high

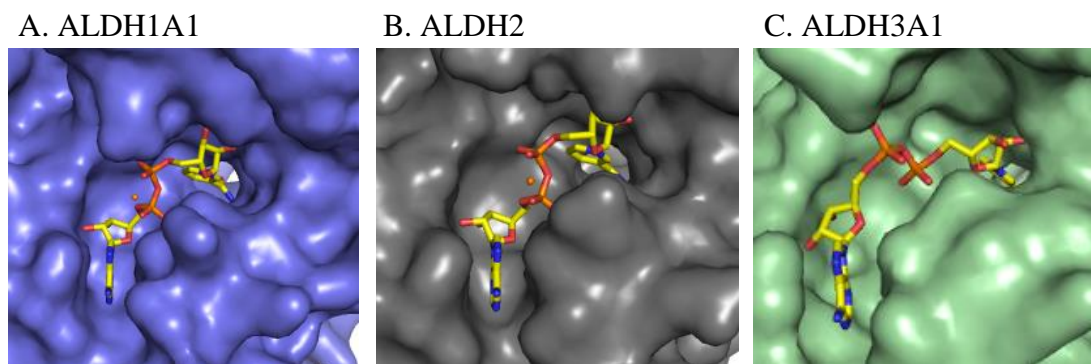


Figure 56: Surface topography of the cofactor binding site for 3 ALDH isoenzymes. A) ALDH1A1 (PDB 4WB9), B) ALDH2 (PDB 1O02), and C) ALDH3A1 (PDB 4L2O). The orange sphere in ALDH1A1 and ALDH2 represent cations, Yb for ALDH1A1 and Mg for ALDH2 that are present during crystallization.

degree of structural similarities, particularly within their coenzyme binding sites (Figure 56). However, there exist distinct differences within their substrate binding sites that could be exploited to discover selective compounds targeting the individual isoenzymes (Figure 12). These structural features support our screening approach to avoid identifying compounds that interact at the cofactor binding site, as they are less likely to be selective for ALDH1/2 class members. It is possible that a number of our esterase activators that had no effect on aldehyde oxidation acted at the cofactor binding site, activating the esterase reaction similar to NAD<sup>+</sup>. However, the near saturating levels of NAD<sup>+</sup> used in the assays would minimize their effect on aldehyde oxidation. If a compound did bind at the cofactor binding site, only an extremely potent or covalent modulator would be identified under these conditions.

The esterase screen used in this study was modeled after a previously reported screen for ALDH3A1 inhibitors that successfully identified two classes of selective ALDH3A1 inhibitors capable of increasing mafosphamide sensitivity in cancer cells<sup>106,136,166</sup>. By adapting this assay to ALDH1A1, we screened a 64,000 compound library and following one round, identified over 900 compounds. Rescreening of these compounds under identical conditions resulted in 256 confirmed hits that modified ALDH1A1 esterase activity. Therefore, the effect on esterase activity of <30% of the identified activators/inhibitors identified in round one were successfully repeated in round two. Although these replicability results may seem low, HTS are inherently noisy and produce many false positives that are eliminated in the second round. As shown in Figure 37, simply calculating the Z-factor produced outliers despite identical conditions within one plate. Some reasons for poor replicability include inaccuracies in compound concentration, spectral interference from the compounds, errors in robot pipetting, and debris or bubbles in the well that interfered with the reading. The second round of screening is designed to remove these false positives from consideration, conserving both time and resources. Since the HTS identified 256 compounds, the large number of false positives from round one was not a concern. In addition, the data can always be revisited at a later date to investigate compounds eliminated in this manner.

We examined the effect on dehydrogenase activity of 67 of the 256 compounds identified via the esterase HTS and found that 30 selectively inhibited ALDH1A1 at least 50% but had little to no effect on ALDH2 or ALDH3A1. One compound inhibited both ALDH1A1 and ALDH2, while a second inhibited ALDH1A1, ALDH2, and ALDH3A1, although ALDH3A1 was only modestly inhibited (40%) versus over 90% inhibition of ALDH1A1 and ALDH2. Therefore, nearly 50% of the esterase modifiers also altered aldehyde oxidation, and most of the compounds did so selectively for ALDH1A1 compared to two other ALDH's. These results indicate that the esterase HTS is well suited for reliably identifying inhibitors of ALDH1A1 aldehyde oxidation activity. However, no confirmed activators of aldehyde oxidation activity have been identified.

Of the 57 esterase activators tested, none activated the aldehyde oxidation reaction of ALDH1A1, but nearly half inhibited it. The esterase reaction is independent of  $\text{NAD}^+$ , but the presence of either  $\text{NAD}^+$  or  $\text{NADH}$  will increase the rate of ester hydrolysis, depending on assay conditions<sup>170,172</sup>. The substrate and cofactor binding sites are linked to the active site by a tunnel through the enzyme. For ester hydrolysis, the substrate can likely enter the active site via either end of this tunnel. To activate esterase activity, it is proposed that cofactor binding slows transit of the ester substrate out of the tunnel, increasing the number of productive encounters with the active site nucleophile and possibly by directly activating the nucleophile (Figure 35C)<sup>140</sup>. Compounds that function as esterase activators but dehydrogenase inhibitors likely bind to the substrate-binding end of this tunnel, as illustrated by Hit-2 in Figure 57. In a manner similar to activation via cofactor binding, compound binding slows the transit of pNPA out of the active site tunnel and increases the likelihood of a productive encounter with the active site cysteine. However, the effect these esterase activators have on the  $\text{NAD}^+$ -dependent aldehyde oxidation reaction is the opposite. Binding of the compound and cofactor alters access to the active site at both ends, and therefore, depending on the structure of the compound and the size of the substrate, could inhibit dehydrogenase activity. However, as seen with the ALDH2 activator Alda-1, a compound that binds at the substrate binding end of the active site tunnel could also result in a dehydrogenase activator, depending on binding position, location relative to the active site residues and substrate size<sup>140</sup>. It is possible

that a number of our esterase activators that had no effect on aldehyde oxidation acted at the cofactor binding site, activating the esterase reaction like  $\text{NAD}^+/\text{NADH}$  (Hit-1 in Figure 57). However, the levels of  $\text{NAD}^+$  used in the assays (approximately  $4 \times K_m$ ) might minimize their effect on aldehyde oxidation, no longer binding to the enzyme in the presence of  $\text{NAD}^+$ , as illustrated by Hit-1 in Figure 57. If a compound did bind at the cofactor site, only an extremely potent or covalent modulator would be identified under these conditions.

Of the 241 esterase activators identified, 78 were structural analogs with a common xanthine ring core structure. Of these 78 compounds, 17 (CM022-031 and CM051-057) were tested further to determine their effect on aldehyde oxidation, with 16 selectively inhibiting ALDH1A1 at least 50% at a compound concentration of 20  $\mu\text{M}$  compared to ALDH1A2, ALDH1A3, ALDH1B1, ALDH2 and ALDH3A1. At 100  $\mu\text{M}$  concentration,

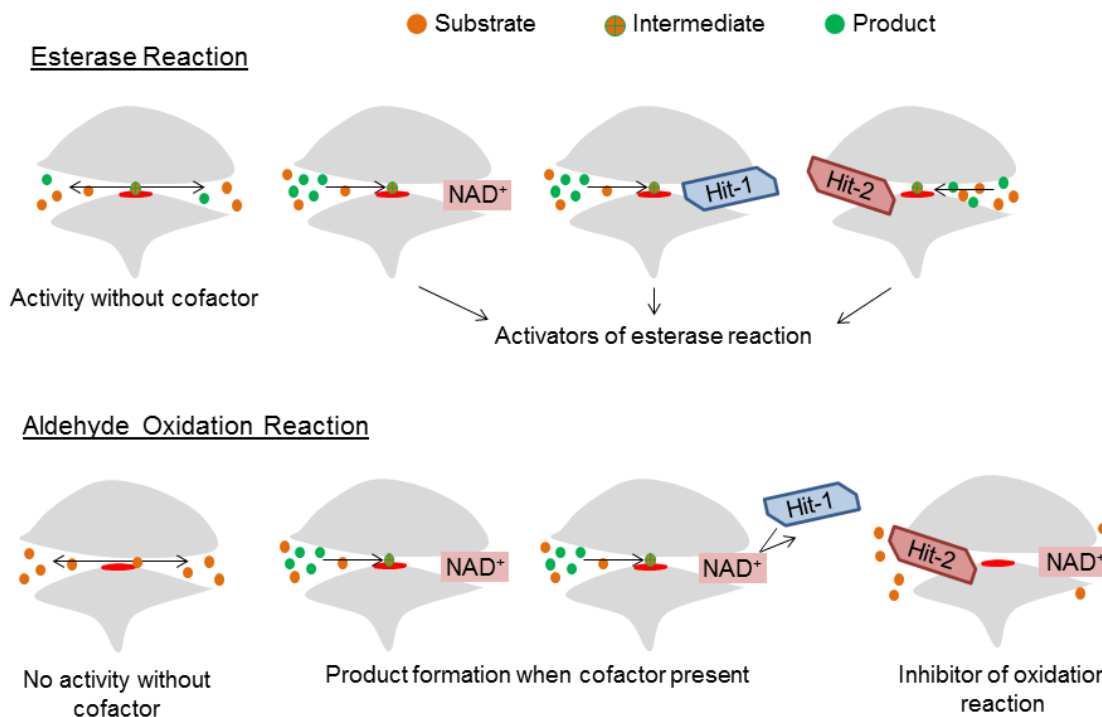


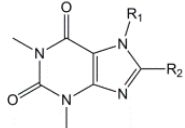
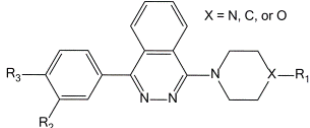
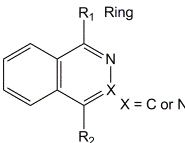
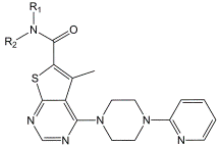
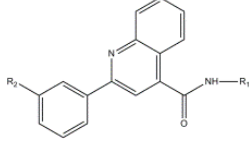
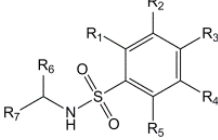
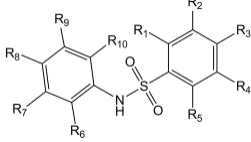
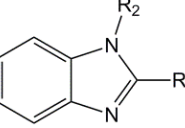
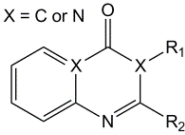
Figure 57: Esterase activators as inhibitors of aldehyde oxidation.

The active site of ALDH1A1 is within a tunnel. For the  $\text{NAD}^+$ -independent esterase reaction, substrate can enter via either side and a compound that binds at either end ( $\text{NAD}^+$ , Hit-1, Hit-2) may possibly activate the esterase reaction. For the  $\text{NAD}^+$ -dependent aldehyde oxidation reaction, Hit-2, a compound that activated esterase reaction could inhibit the oxidation reaction by blocking access to the active site.



some CM026 analogs activated other members of the ALDH1/2 family, mostly ALDH1A3, although with just a single dose, it is not known whether these results represent enzyme activation or something else is occurring to increase absorbance at 340 nm. The esterase HTS also produced eight other structural groups that contained seven or more analogs. As a consequence, 65% of the HTS esterase hits could be classified into nine structural categories (Table 11). An additional eight structural groups contained 2 – 6 analogs each plus 22 structurally unique compounds. As the majority of esterase activators, including a number of structurally unique compounds and classes of compounds, have yet to be tested, it is probable that at least one ALDH1A1 aldehyde oxidation activator is present. An analog of Alda-1 (CM021) showed modest activation of the mutant ALDH1A1 N121S, but did not activate WT enzyme. Of the 15 esterase inhibitors identified in the HTS, only two were structurally similar. The effect on aldehyde oxidation of ten esterase inhibitors has been further examined, with four inhibiting ALDH1A1 at least 50% at a concentration of 20  $\mu$ M. CM302 was a nonselective inhibitor, inhibiting ALDH1A1, ALDH2, and ALDH3A1, but it was a particularly potent inhibitor for the ALDH1/2 family. CM303 inhibited ALDH1A1 and ALDH2 but not ALDH3A1, while both CM306 and CM307 inhibited ALDH1A1 but not ALDH2 nor ALDH3A1. Additional work is needed to determine the effect of CM303, CM306, and CM307 on other members of the ALDH1 family.

Table 11: Structural classes of hit compounds.

Structure	HTS Hits	Dehydrogenase Activity	
		Tested	Results
	78	17	16 Inhibitors
	20	3	CM001
	10	1	No effect
	9	0	
	13	1	No effect
	11	4	3 Inhibitors
	8	3	CM047
	7	4	CM010
	7	0	

Results of nine structural classes of esterase modulators representing 65% of the compounds identified from the esterase HTS.

Results from kinetic studies indicated that CM026 displayed a noncompetitive partial mode of inhibition with respect to varied substrate acetaldehyde and an uncompetitive partial mode of inhibition with respect to the cofactor  $\text{NAD}^+$ . In a classical one-substrate system, a noncompetitive inhibitor binds to both the free enzyme and the enzyme-substrate species, forming the enzyme-substrate-inhibitor (ESI) complex. An uncompetitive inhibitor binds only to the enzyme-substrate species, forming the ESI complex. With complete inhibition, the ESI complex is inactive, while partial inhibition indicates the ESI complex retains activity, although at a decreased rate of product formation. However, ALDH1A1 aldehyde oxidation is a two-substrate system; the aldehyde itself and the cofactor  $\text{NAD}^+$  are both considered substrates. A schematic of the steady-state kinetic results is shown in Figure 58. The kinetic results indicate that CM026 binds to the enzyme- $\text{NAD}^+$  and enzyme- $\text{NADH}$  complexes. It may be possible that the enzyme- $\text{NAD}^+$ -inhibitor complex is able to complete a cycle. As a partial inhibitor retaining approximately 15% the activity of uninhibited enzyme, either the enzyme- $\text{NAD}^+$ -substrate-CM026 complex and/or the binding of CM026 to the enzyme- $\text{NADH}$

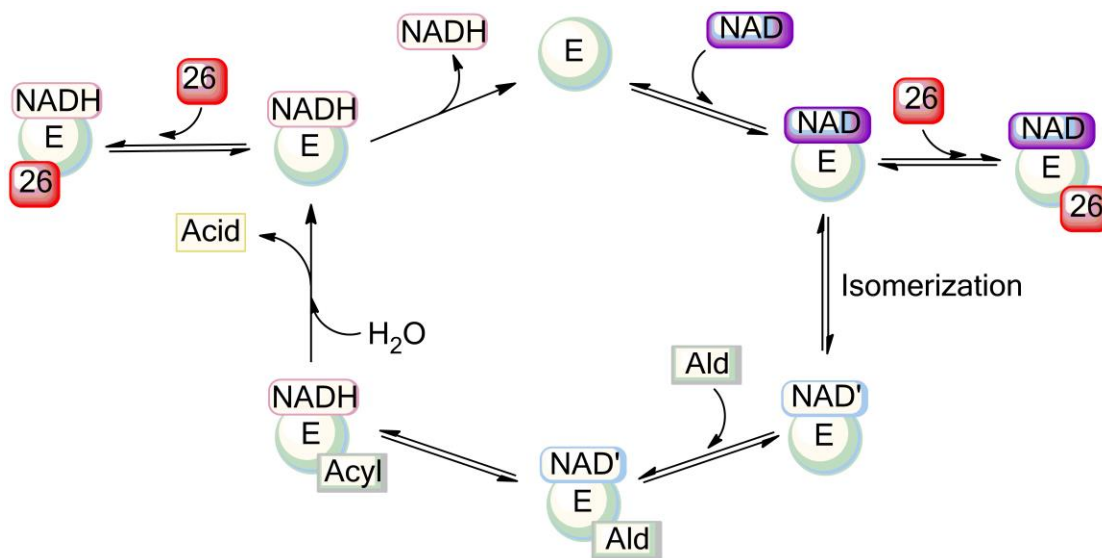


Figure 58: Schematic representation of CM026 inhibition of ALDH1A1. CM026 was noncompetitive with respect to varied acetaldehyde with saturating  $\text{NAD}^+$  and uncompetitive with respect to varied  $\text{NAD}^+$  and saturating propionaldehyde indicating that CM026 can bind to either the enzyme- $\text{NAD}$  complex or the enzyme- $\text{NADH}$  complex. As a partial inhibitor, the enzyme remains active in the presence of CM026. However, the rate of product formation is much slower.

complex has much lower activity than control. In the presence of CM026, the order of the release of the products and inhibitor has not been determined, but it is probable that the acid product leaves first followed by either CM026 or the reduced cofactor. Although kinetics indicate that CM026 only binds to certain enzyme-NAD(H) species, crystallographic results using high concentrations of the compound show it can bind to the free enzyme. The more potent CM026 analog, CM053 displayed noncompetitive tight binding inhibition compared to varied substrate acetaldehyde. In a reaction, an inhibitor can exist in two general phases: free inhibitor ( $I_{\text{free}}$ ) and inhibitor bound to an enzyme (EI), with multiple species of the EI complex possible. In standard inhibition calculations, the amount of inhibitor (I) greatly surpasses the amount of enzyme (E), and therefore when determining the total amount of I ( $I_{\text{total}}$ ) in the system, EI can be disregarded because it is so small in comparison to  $I_{\text{free}}$ . However, for inhibitors that bind tightly to the enzyme, steady-state kinetics will require much lower concentrations of the inhibitor and a significant portion of inhibitor will be found bound to the enzyme. Therefore, EI cannot be disregarded in the calculations. For CM053, the mode of inhibition kinetics used 150 nM enzyme and a maximum of 300 nM inhibitor. Unlike CM026, binding of CM053 will ultimately produce an inactive enzyme. CM028 contains a larger side chain at R1 compared to CM053, replacing an isopentyl group with a phenylpropyl group and exhibits competitive tight inhibition with respect to acetaldehyde.

There were 77 compounds in the initial screening results that were structurally similar to CM026. We tested 17 members of this compound class. Unlike traditional SAR, this SAR only contained compounds known to bind to ALDH1A1, as they all modified the enzyme's esterase activity (Table 8). Halogens on either R-group were not well tolerated and only less potent inhibitors possessed a halogen. This is particularly obvious with the R1 group as it projects into a hydrophobic pocket formed by two phenylalanines (F171 and F466). Less clear with the R1 group is why a linear butyl group was not favored but a branching isopentyl group was found in 3 of the 4 most potent inhibitors, based on their  $IC_{50}$  values. For potent inhibition of aldehyde oxidation, the R2 group must have hydrogen bonding capability. The three compounds with sub-micromolar  $IC_{50}$  values contain amide groups, with the carbonyl carbon six atoms removed from the xanthine

ring. CM055 and CM026 are the second and third most potent analogs, respectively, and they share a common R2 group. They differ in their R1 group, with CM026 possessing an isopentyl group while CM055 has a larger hydrophobic 3-methylbenzyl group, which may bind tighter in the R1 hydrophobic pocket and decrease the  $IC_{50}$  value nearly 4-fold. However, there is a limit to the size of the hydrophobic group on R1. CM053 and CM028 share the same R2 group but CM053, the most potent inhibitor, has an isopentyl group at R1 while CM028 has a phenylpropyl group. This difference in the R1 group results in a nearly 10-fold increase in  $IC_{50}$  values between CM053 and CM028. It is possible that the methylbenzyl group at R1 (seen in CM055) combined with the 1-methylpiperidine 4-carboxamide group at R2 (seen in CM053) would be a more potent inhibitor than CM053, but no compounds with a similar structure were included in the HTS library.

CM037 was a competitive tight inhibitor with respect to varied acetaldehyde, with a  $K_i = 230$  nM, and uncompetitive with respect to  $NAD^+$ . The schematic for ALDH1A1 inhibition by CM037 is shown in Figure 59. Unlike CM026 and CM053, there is no crystallographic evidence to indicate that CM037 binds to the free enzyme. As a competitive inhibitor, binding of CM037 and acetaldehyde is mutually exclusive,

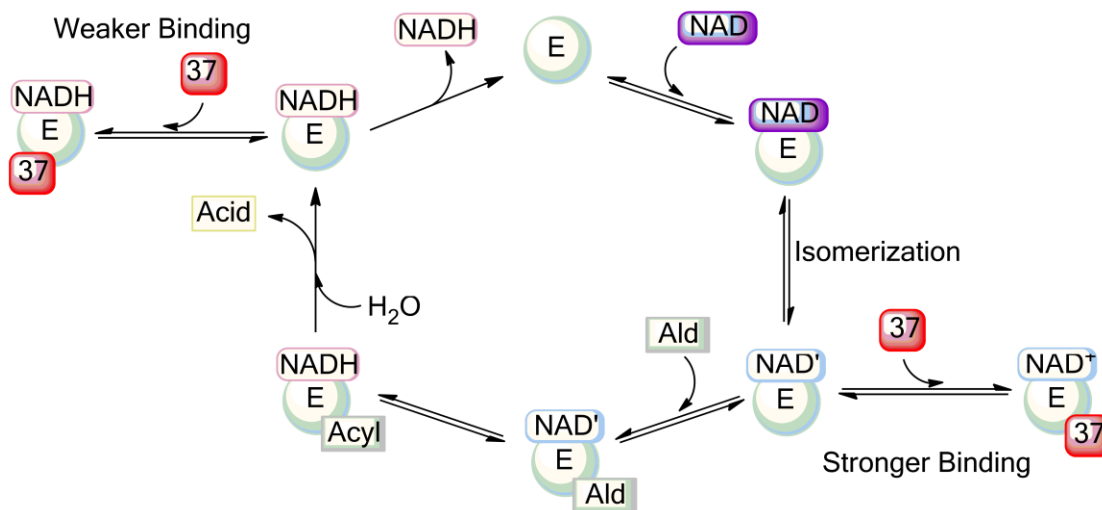


Figure 59: Schematic representation of CM037 inhibition of ALDH1A1  
 CM037 is a competitive inhibitor with respect to acetaldehyde and uncompetitive with respect to  $NAD^+$ . The predominant pathway is CM037 binding to the enzyme- $NAD'$  complex. However, inhibition is still present with saturating aldehydes, and therefore CM037 can likely bind to the enzyme- $NADH$  complex.

resulting in an inactive enzyme. However, even in the presence of saturating aldehyde, CM037 still inhibits to some extent. As seen with CM026, CM037 was uncompetitive with respect to varied  $\text{NAD}^+$  and likely binds weakly to the enzyme-NADH complex, as the  $K_i$  value with respect to varied  $\text{NAD}^+$  and saturating propionaldehyde was at least 20-fold weaker than with respect to varied acetaldehyde and saturating  $\text{NAD}^+$ . CM037 is an ester and it is possible that it is a substrate for ALDH1A1 (Figure 60). In a manner similar to inhibition of ALDH1A1 by DEAB, this ester may be turned over at a slow rate, therefore appearing as an inhibitor. It is also possible that the carboxylic acid product acts as an inhibitor, with either the product remaining bound to the enzyme (slow turnover rate) or the product released only to rebind and inhibit the enzyme, although preliminary results indicate the acid product does not inhibit the enzyme. However, structural data shows that the ester is still intact after a five hour incubation with  $\text{NAD}^+$ , an activator of esterase activity. Also, structural data does not indicate that CM037 is binding in an orientation that would support a substrate as inhibitor mechanism. SAR with CM037 indicates that the ester functional group is important for ALDH1A1 inhibition. Replacing the ester group with an amide resulted in a 5-fold increase in  $\text{IC}_{50}$  value, suggesting that CM037 may inhibit by acting as a substrate. The only non-ester analog that inhibited ALDH1A1 at a similar level to CM037 was the carboxylic acid CM037e. Although the carboxylic acid functional group of CM037e is on R1, whereas the carboxylic acid functional group for the CM037 product would be on R2, these results suggest that the esterase product may be able to bind to and inhibit ALDH1A1. The substituent on R1 is important for isoenzyme selectivity within the ALDH1A subfamily. CM037 does not inhibit ALDH1A2 or ALDH1A3 more than 20% at a concentration of 20  $\mu\text{M}$ . However,

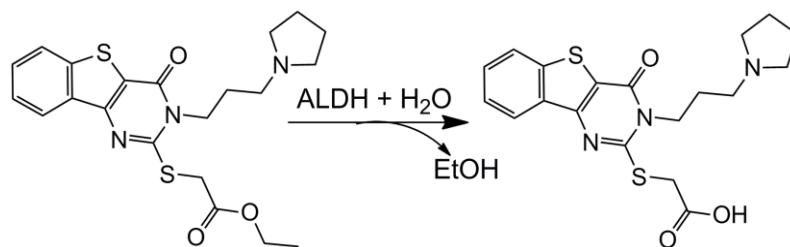


Figure 60: Possible hydrolysis of CM037 by ALDHs. CM037 is an ester that possibly acts as a substrate for ALDH1A1, producing ethanol and a carboxylic acid.

replacing the propyl pyrrolidine group with either an ethanol group (CM037g) or a phenyl group one or two carbons removed from the core three-ring structure (CM037b, CM037j), greatly increased ALDH1A2 and ALDH1A3 inhibition. CM037b, CM037c and CM037j, which all contain phenyl groups on R1, were not strong inhibitors of ALDH1A1. Either a longer, more flexible linker is required for binding or the aromatic ring interferes with compound binding.

As shown in Figure 61, ALDH1A1 possesses a large, funnel shaped opening leading to the active site, whereas ALDH2 has a much more constricted, cylindrical shaped opening. ALDH3A1 possesses a wider inner vestibule near the catalytic nucleophile, with a much narrower and curved entryway. The cylindrical entry to ALDH2 sterically occludes the binding of the CM026 class of inhibitors in large part due to the side chain present at the position equivalent to Gly458. Like daidzin, both the CM026 and CM037 classes of compounds are planar, multi-ringed structures that adopt binding modes that take advantage of the topological characteristics unique to the ALDH isoenzyme toward

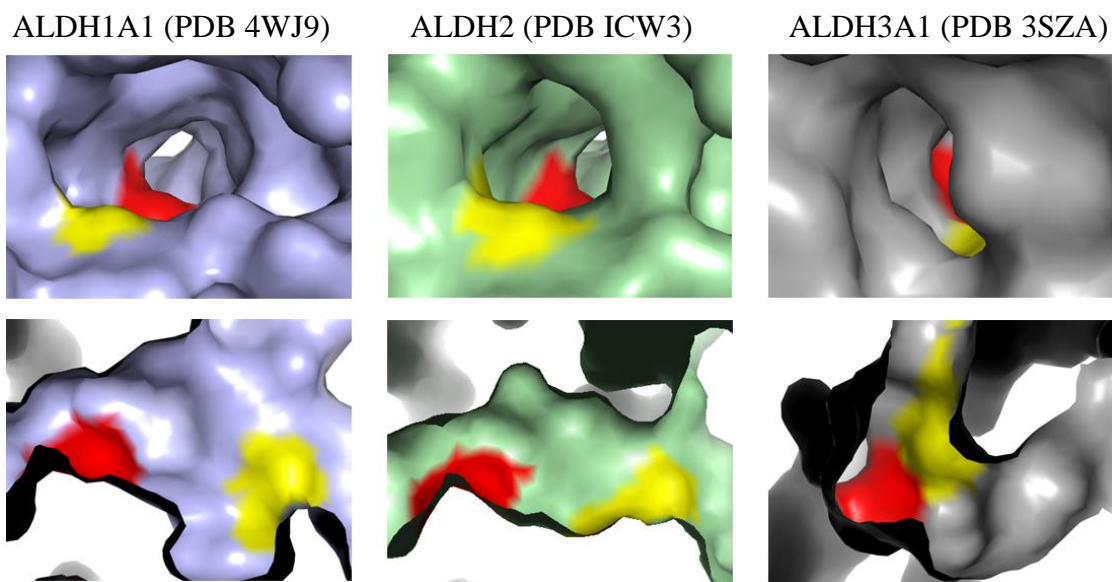


Figure 61: Comparison of the active site topography of human ALDHs.

The three isoenzymes were aligned using LSQ in Coot and the surface figures generated via Pymol. The active site cysteine is shown in red for all three isoenzymes. In ALDH1A1, Gly458 is shown in yellow. For ALDH2 and ALDH3A1, the amino acid at the equivalent position as ALDH1A1 G458 is also shown in yellow (ALDH2 Asp457, ALDH3A1 Ile394).

which they demonstrate selectivity, and neither of these types of compounds can be accommodated in the curved and more sterically restricted substrate binding site of ALDH3A1. Daidzin is a strong inhibitor of ALDH2 but it also inhibits ALDH1A1<sup>147</sup>. Comparison of the binding of daidzin to ALDH2 (PDB Code 2VLE<sup>173</sup>) with CM026 binding to ALDH1A1 indicate that they bind in a similar location but in a different orientation. This near perpendicular difference in orientation enables CM026 to inhibit ALDH1A1 but not ALDH2 due to steric hindrance of Asp474 of ALDH2, while daidzin inhibits both isoenzymes (Figure 62A). Although CM026 and CM037 have little compound similarity, they utilize the same regions of the active site of ALDH1A1, with both capitalizing on the presence of the small binding pocket created by the presence of Gly458 (Figure 62B). Sequence alignment of the other 18 human ALDH isoenzymes indicate that the only other family member that has a glycine in this position is ALDH16A1. However, the function of ALDH16A1 is unknown, and the protein has a three residue deletion in the active site loop that eliminates the conserved Cys nucleophile, suggesting this protein may lack aldehyde dehydrogenase activity<sup>36</sup>. Analysis of the G458N mutant indicates that the presence of a non-glycine residue at this

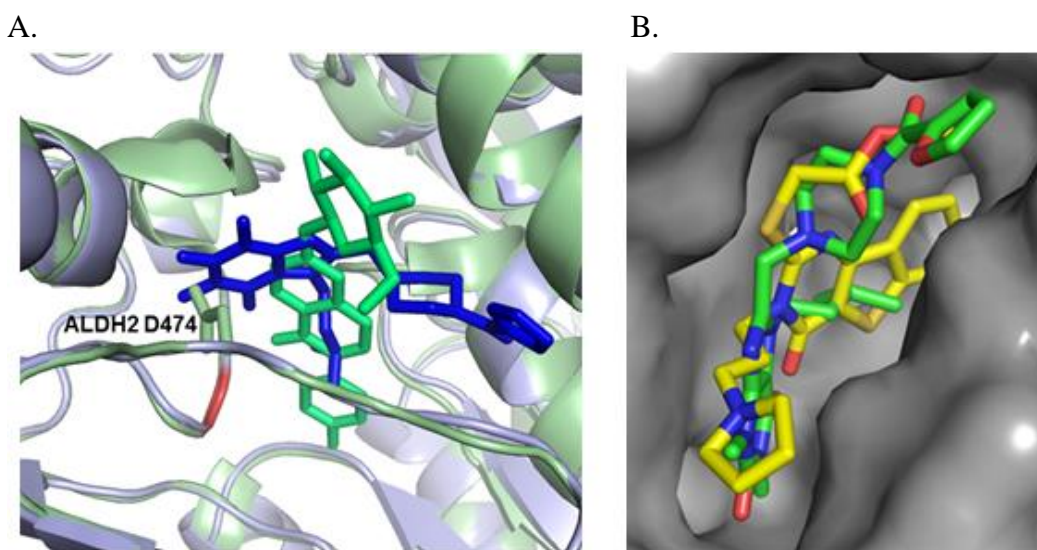


Figure 62: Comparison of compound binding to ALDH isoenzymes. A) Comparison of ALDH1A1-CM026 to ALDH2-daidzin with ALDH1A1, shown in light blue with CM026 in dark blue, and ALDH2, shown in light green with daidzin in dark green. The active site cysteine is shown in red for both isoenzymes. B) Comparison of CM026 and CM037 binding to ALDH1A1 illustrating that both classes utilize similar surface features of ALDH1A1. Daidzin-ALDH2 (PDB 2VLE).



position does not adversely affect either substrate kinetics or inhibition by the non-selective compounds DEAB and CM302. Consequently, our screening protocol successfully exploited a unique structural feature found primarily in primate ALDH1A1s, which further validates the use of the esterase activity as a screening tool for ALDH isoenzymes.

Unlike DEAB, both the CM026 and CM037 classes of compounds were selective for ALDH1A1, and our next question was to determine if they can produce a physiological effect in cells. For a compound to have an effect in the cell, it must have certain characteristics in order to reach its protein destination. First, it must be able to pass through the cell membrane and enter the cell. Next, the compound must not be inactivated or exported from the cell by active transporters too quickly. Third, it must modulate the activity of its target with minimal off-target effects. Finally, the compound itself or any of its metabolic by-products must not exceed toxicity thresholds<sup>174</sup>. If any or all of these characteristics are not present, even the most potent *in vitro* modulators will fail in cell culture.

The next step in the development of the CM026 and CM037 classes of compounds was to determine whether these compounds could produce a desired biological effect in cell culture. For this step, we recruited two cancer researchers: Dr. Daniela Matei at Indiana University School of Medicine and Dr. Ronald Buckanovich at the University of Michigan School of Medicine. In both laboratories, the effects of the CM026 and CM037 classes of compounds on ovarian cancer (OC) cell models were tested. OC cells form cellular aggregates, or spheroids, that facilitate metastasis by serving as vehicles for dissemination, while protecting the cells from the extracellular, peritoneal environment<sup>175,176</sup>. Dr. Matei has shown that ALDH1A1 is over-expressed in spheroids compared to cellular monolayers using both OC cell lines and primary cells, and that ALDH1A1 mRNA and protein levels increased with each successive spheroid generation<sup>101</sup>. These results indicate that the protein is not simply a biomarker but is important for cellular aggregation and/or cancer stem cell survivability. In a dose-dependent manner, the Matei group showed that CM037 was able to disrupt spheroid

formation and decrease cell viability by targeting ALDH<sup>+</sup> cell populations (Figure 63). Via the Wnt signaling pathway, the ALDH1A1 gene is a direct target of  $\beta$ -catenin, a transcription factor linked to stem cell self-renewal, survivability, and as discovered by Matei *et al*, the formation of OC spheroids. Following inhibition of ALDH1A1, the downstream effects of ALDH1A1 inhibition that disrupt spheroid formation are not known, as the role of this enzyme in stem cell maintenance, self-renewal, and/or survivability in general are still unclear. CM037 and other selective small molecule inhibitors of ALDH1A1 should increase our understanding of these processes. ALDHs are also involved in increasing cellular resistance to various chemotherapeutics, and CM037 increased cellular sensitivity to cisplatin in a dose-dependent manner<sup>101</sup>. These results with OC cells indicate that CM037 is capable of entering a cell and inducing a phenotypic change. Therefore, CM037 represents a potent first-generation selective inhibitor of ALDH1A1 that should be further developed and modified to improve potency and selectivity.

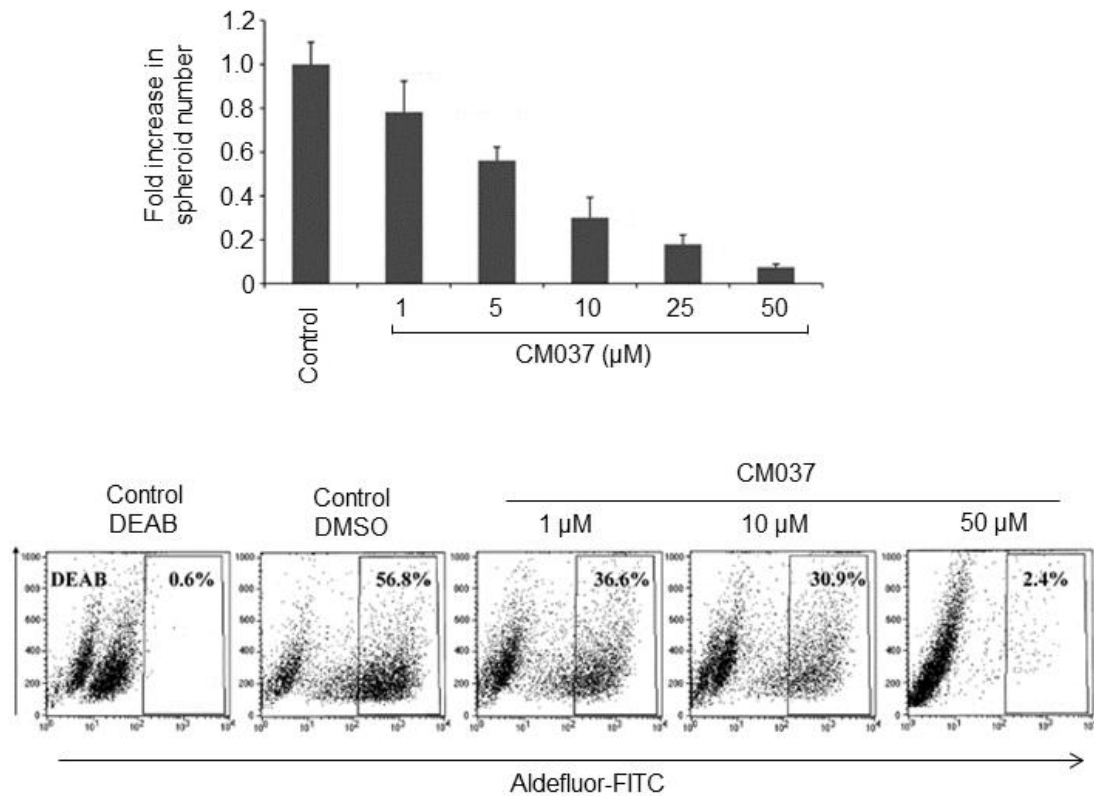


Figure 63: The effect of CM037 in ovarian cancer cells. Dose-dependent decrease following 3-day CM037 treatment in A) spheroid formation and B) ALDH<sup>+</sup> cells (modified from Condello *et al*<sup>101</sup>).

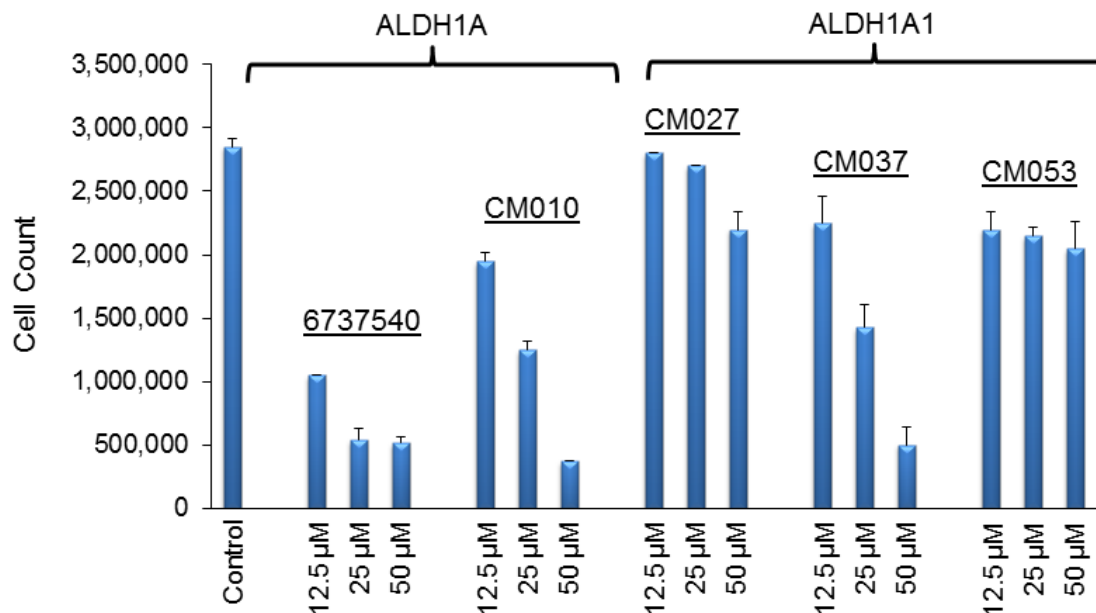


Figure 64: Effect of various compounds on the proliferation of OC cell. A2780DK, an ovarian cancer cell line, was plated with an initial cell count of 500,000 cells. Following a 2-day treatment with ALDH1 inhibitors, cellular proliferation was determined by counting the number of viable cells. Compounds 6737540 and CM010 inhibit the ALDH1A subfamily, and CM027, CM037, and CM053 are selective for ALDH1A1. 6737540 is a University of Michigan compound developed by Ronald Buckanovich and used here as a control (bars represent mean/SEM; unpublished data).

The CM026 class of compounds has been less successful in cell culture. Tested by the Matei lab along with CM037, CM026 did not disrupt spheroid formation. In unpublished work from the Buckanovich lab, neither CM027 nor CM053, two analogs of CM026, altered cellular proliferation in OC cells, while CM037 did so in a dose dependent fashion (Figure 64). Why the CM026 analogs are ineffective in cell culture are not known. The CM026 analogs tested had low partition coefficients (logP) values, an estimate of the compound's lipophilicity, and therefore its cell membrane permeability. All compounds had logP values less than 5, in accordance with Lipinski's Rule of 5<sup>151</sup>. However, the three CM026 analogs tested in cell culture, CM026, CM027, and CM053 had estimated logP values of 1.99, 2.25, and 1.75, respectively, based on ChemDiv calculations ([www.chemdiv.com](http://www.chemdiv.com)). These values indicate that the compounds may be too hydrophilic to be efficiently taken up by the cell. Of the CM026 class of compounds tested, CM057 had an  $IC_{50} < 1 \mu M$  and an estimated logP of 3.44, and may serve as a

good representative of this class in cell culture. However, maximum solubility in DMSO is approximately 10 mM, and CM057 was not tested in cell culture due to limitations on maximum compound concentration. It may be possible to use a different solvent to achieve higher concentrations for CM057. Another approach is to test the remaining CM026 analogs identified in the HTS that have higher logP values. The logP values of this class ranged from 1.71 to 4.92, so ample variability exists to determine whether the inability of the CM026 class of compounds to alter cellular phenotype is simply due to the compounds not getting into the cells. In comparison to CM026, CM037's logP value was 4.02 and likely could easily enter the cell. However, CM037 solubility has become an issue moving forward from a cell culture model to an animal model, as it is far less soluble in the required aqueous solutions. Other possible reasons for why the CM026 analogs have no effect in cell culture include: 1. the compound is taken up by the cell but it is quickly metabolized into an inactive form, 2. the compound is taken up by the cell but is actively transported back out of the cell, 3. the compound binds to another macromolecule and is not available to inhibit ALDH1A1, and 4. the compound is ineffective against ALDH1A1 in a cellular environment. Additional cell culture studies are needed to determine why this potent class of ALDH1A1 inhibitors is ineffective in the cellular environment. Although cell culture data with the CM037 class is much more promising, the CM026 class cannot be eliminated. The results from the Buckanovich lab also revealed that CM010, another compound discovered in the HTS, reduces cell proliferation in a dose-dependent manner (Figure 64). CM010 is not selective for ALDH1A1, inhibiting the other two ALDH1A subfamily members, but has no effect on ALDH2 or ALDH3A1. As CM010 was more potent in this model compared to CM037, it may either be a more effective inhibitor of ALDH1A1 in cells or the combined inhibition of multiple ALDH1A isoenzymes by CM010 was more effective in blocking proliferation of the cancer cells. Therefore, all three classes of compounds (CM010, CM026, and CM037) still remain excellent starting points for the development of ALDH1A and ALDH1A1 selective compounds.

## V. Conclusion and Future Directions

Members of the aldehyde dehydrogenase superfamily of enzymes are critical for normal biological processes. ALDH1A1 has also been linked to a wide variety of diseases including obesity, cancer, and neurodegeneration. Small molecule probes are needed to study the role of this enzyme in both normal and disease states, but due to the high degree of structural and functional overlap within the ALDH1/2 family, targeted drug development has proven to be difficult. At this time, there are no commercially available small molecules that selectively modulate ALDH1A1 activity compared to other ALDH isoenzymes. The goal of our study was to discover and characterize ALDH1A1-specific inhibitors.

Our results with DEAB indicate that the substrate-as-inhibitor type of compounds may not represent good candidates for the development of selective ALDH inhibitors due to the high degree of substrate overlap amongst ALDH superfamily members, particularly within the ALDH1/2 family. A number of ALDH superfamily members are capable of the oxidation of DEAB to its corresponding carboxylic acid, but with great variation in the rate of the reaction. DEAB is a potent inhibitor of ALDH1A1, but it also inhibits most of the ALDH1/2 family members and in the process is metabolized to its carboxylic acid and removed from the system, allowing for short-term, reversible inhibition. This overlap of inhibition may be of minimum concern if a short-term knockdown of global ALDH1/2 activity is desired. However, DEAB is rapidly turned over by ALDH3A1, and the presence of even modest levels of this enzyme could hinder the use of DEAB in a system or tissue-type with high levels of ALDH3A1 expression. Analogs of DEAB could be developed that are not metabolized as rapidly by ALDH3A1. To better design these analogs, the structures of DEAB with multiple ALDH isoenzymes are needed to determine how the compound binds and which residues are involved in stabilizing the acyl-enzyme intermediate. As ALDH3A1 prefers aromatic substrates like DEAB and benzaldehyde compared to the linear aliphatic aldehydes preferred by the ALDH1/2 family, structural data could facilitate the development of DEAB analogs that are no longer substrates for ALDH3A1. However, due to the high level of structural and functional overlap within the ALDH1/2 family, it will likely be very difficult to develop

DEAB analogs that are selective towards one isoenzyme. DEAB is a potent, reversible ALDH1/2 inhibitor, but it is not a valid candidate for the development of inhibitors selective for a particular isoenzyme. Therefore, new classes of compounds are needed.

To discover selective activators and inhibitors for ALDH1A1, an *in vitro* esterase-based high-throughput screen (HTS) was performed. From a 64,000 compound library, 256 compounds were identified as hits capable of altering ALDH1A1's esterase activity. We have further examined 67 compounds and discovered approximately 50% also inhibited aldehyde oxidation. These results indicate that the esterase screen is a valid option for identifying ALDH1A1 aldehyde oxidation modulators.

We further characterized two classes of compounds from the HTS in detail, CM026 and CM037, using steady-state kinetics, structure-activity-relationship, X-ray crystallographic structural data, and cell culture studies. Currently CM037 is the most promising and is capable of inducing a phenotypic change in two ovarian cancer cell lines (Condello *et al*<sup>101</sup> and Buchanovich *et al*, unpublished). As CM037 progresses from hit to lead compound, additional pharmacokinetic studies are needed to determine if CM037 can become an effective drug. These studies will address CM037's absorption, distribution, metabolism, and excretion (ADME) properties<sup>174</sup>. CM037 is an ester and even if CM037 is not a substrate of ALDH1A1, other esterases exist in the cell that will likely hydrolyze the compound. Microsomal stability assays will aid in better understanding the clearance rate of the compound. Toxicity studies in mice are in the early stages by the Matei lab. One problem with these toxicity studies has been solubility issues of CM037 in an aqueous solution, and therefore DMSO has been needed at a higher percentage than desired, causing temporary discomfort to the mice. With doses at 30 mg/kg or below, there have been no apparent effects on the mice. Additional pharmacokinetic and pharmacodynamics studies are needed in mice to determine the LD<sub>50</sub> (dose required to kill one-half of the animals), the rate of compound elimination from the body, optimal route of administration, and other drug ADME characteristics. CM037 has an *in vitro* IC<sub>50</sub> value of approximately 5 μM, and cell culture results had EC<sub>50</sub> values in the 20 – 50 μM range. In its current form, CM037 does not represent an ideal drug, but modifications

could increase potency and improve solubility while maintaining selectivity for ALDH1A1. Therefore, CM037 represents an excellent choice for further drug development.

Since CM026 and its analogs represented nearly 25% of both the 256 hits identified in the esterase HTS and the 67 compounds analyzed further for their effect on aldehyde oxidation, this class of compounds was studied the most. Kinetic studies showed that this class contains potent, selective, reversible ALDH1A1 inhibitors. Structural data enabled us to determine where the compounds bind and why they are selective for ALDH1A1 compared to other ALDH1/2 family members. Unfortunately, cell culture data was less promising, as the compounds selected for cell culture studies had little to no effect on cellular phenotype, viability, or proliferation. However, it is possible that the compounds have little effect on cells because they are too hydrophilic and simply are not getting into the cells. Of the 17 CM026 analogs tested, CM057 represents the most promising compound for possible cell culture studies. CM057 has a sub-micromolar  $IC_{50}$  and an estimated  $\log P$  value of 3.44, similar to CM037 and is more hydrophobic than the other CM026 analogs tested in cell culture. Because CM057 is less soluble in DMSO than the other CM026 analogs, it was not explored further, but this characteristic may make it more promising for future *in vivo* testing. If CM057 solubility issues continue to be problematic, there exists a great deal of diversity amongst the nearly 75 CM026 analogs identified in the HTS, and compounds that both inhibit ALDH1A1 aldehyde oxidation and are cell permeable likely exist. Structural data, SAR results, and estimated  $\log P$  values will help determine which hits represent the best candidates for further development of the CM026 class of compounds. Unfortunately, a negative result still does not confirm that the compound is ineffective in cell culture. Solubility or permeability issues, inactivation of the compound, and/or rapid removal of the compound from the cell will also produce negative results. Regardless, by testing a variety of CM026 analogs of varying  $\log P$  values, we should be able to determine if this class of compounds represents a potential candidate for the development of selective ALDH1A1 inhibitors.

The CM026 and CM037 classes of compounds represented 18 of the 67 compounds further analyzed and 13 of the 32 hits that inhibited ALDH1A1, leaving 19 compounds yet to be analyzed in detail that are known to inhibit ALDH1A1. CM302 and CM303 inhibited both ALDH1A1 esterase and dehydrogenase activities but were not selective for ALDH1A1 compared to other ALDH isoenzymes tested. However, if a short term inhibition of ALDH1/2 activity, irrespective of any specific isoenzyme, is desired, both of these compounds represent new classes of compounds with this capability. Unlike DEAB, CM302 has an added feature in that it also partially inhibits ALDH3A1 and possibly ALDH5A1, and may represent an ideal candidate for drug development targeting cells that overexpress both ALDH1A1 and ALDH3A1. The esterase activator CM010 is a selective inhibitor of the ALDH1A subfamily, and results from the Buchanovich lab at the University of Michigan indicate that this compound decreases cell proliferation. Similar to CM037, CM010's *in vivo* EC<sub>50</sub> was in the 20 – 50 μM range with an *in vitro* IC<sub>50</sub> in the low micromolar range, but unlike CM037, complete inhibition was not seen under the assay conditions used. Additional modifications are needed to improve the potency of CM010. Very little has been done on the remaining 16 compounds, but all are selective for ALDH1A1 compared to ALDH2 and ALDH3A1. The most promising of these unexplored compounds are CM020, CM038, CM039, CM047, and CM307, as all have sub-micromolar IC<sub>50</sub>'s. Although CM039 has some structural resemblance to CM037, the remaining compounds are all structurally unique, and each could represent a new class of ALDH1A1 inhibitors. Further kinetic and cell culture work is needed to determine if any or all of these compounds should be pursued as potential hits. Additional work is also needed to characterize all classes of compounds identified in the HTS. Although most of the 15 esterase inhibitors have been examined, with 4 of 10 also inhibiting aldehyde oxidation, much more work is needed on the esterase activators. Based on the results of the HTS compounds tested to date, almost half of these activators will likely modify ALDH1A1 dehydrogenase activity, and most will inhibit the enzyme selectively, generating even more candidates for the development of selective ALDH1A1 inhibitors.



The HTS identified 241 compounds that activated ALDH1A1 esterase activity, but upon examination of 57 of these activators, none activated the enzyme's aldehyde oxidation activity. One compound, CM021 exhibited a modest increase of ALDH1A1 N121S activity, but did not activate ALDH1A1 WT. CM021 is a structural analog to Alda-1, an ALDH2 activator identified by coupling aldehyde oxidation to an NADH-dependent reaction<sup>75</sup>, indicating that the HTS is capable of identifying activators. Although an activator of aldehyde oxidation was not in the first 57 esterase activators analyzed, it is still probable that at least one exists within the compounds remaining to be analyzed.

Our study has identified and characterized multiple classes of compounds that inhibit ALDH1A1, including some that exhibit promising anti-cancer properties in cell culture studies. ALDH1A1 inhibition likely induces this phenotype by changing the levels of various retinoids, and therefore alters various cell growth and differentiation pathways. Although our goal was to develop selective ALDH1A1 compounds, we have also discovered and characterized compounds that inhibit the ALDH1A subfamily and multiple members of the ALDH1/2 family. Each compound class could serve as a scaffold to design drugs that target the ALDH1/2 family of enzymes, the ALDH1A subfamily, or ALDH1A1 specifically. For inhibition of the ALDH1/2 family, reversible inhibition by CM302 and CM303 would be preferred to minimize the long-term consequences of such a global reduction in enzyme activity, including the build-up of toxic aldehydes that would occur with an irreversible inhibitor. Short-term, reversible inhibition of ALDH1A activity, as seen with CM010, may be ideal to minimize the cancer stem cell characteristics identified as contributing to more aggressive and fatal cancers. In combination with other drugs, inhibition of ALDH1A may prevent CSC proliferation and enable complete destruction of the cancer. Long-term inhibition of ALDH1A may not be desired due to the critical role of this subfamily in a number of cell growth and differentiation pathways. Finally, for selective inhibition of ALDH1A1, both the CM026 and CM037 classes of compounds can serve as a scaffold for the continued development of drugs targeting this enzyme. Initially, these ALDH1A1-selective compounds would be of great use as chemical tools for deciphering the role of this enzyme in multiple biological pathways, both in normal and disease states. With the use

of various “-omics”, including metabolomics and proteomics, as well as other biological tools, it is possible to decipher the downstream effects of ALDH1A1 activity on various biological processes. By increasing our knowledge of the processes involved, we can better understand the physiological roles of ALDH1A1 and determine which steps to target for drug development. If ALDH1A1 is discovered to be a potential target for chemotherapeutics, these selective ALDH1A1 inhibitors again represent a potential starting point for drug development. By precisely targeting ALDH1A1, the development of more potent, longer lasting covalent inhibitors may be possible, as other similar ALDH isoenzymes could minimize the build-up of toxic aldehydes. Regardless of the degree of selectivity desired, from ALDH1A1-selective to a global inhibition of multiple ALDH enzymes, the compounds identified in this study represent an excellent starting point for the development of ALDH1A1 inhibitors.

Supplementary Table 1: Activators of ALDH1A1 esterase activity.

ChemDiv catalog number and structure of the 241 compounds that activated ALDH1A1 esterase activity in the HTS. The ratio to control value for the initial screen (1) and verification screen (2) are shown. The names of the compounds whose effect on aldehyde oxidation have been tested are shown as CM0XX below the ChemDiv number.

Chem Div	Structure	Ratio to Control		Chem Div	Structure	Ratio to Control	
		1	2			1	2
2188-3777		2.1	2.9	3339-3595		7.3	4.0
3238-0138 CM052		7.5	5.1	3473-1281		4.8	4.9
3238-0120		4.5	4.3	3544-1012 CM003		2.4	2.3
3277-0112 CM046		4.0	2.4	3544-1018		2.4	2.0
3335-0338		3.2	2.2	3635-1425		2.7	2.1
3332-1185		4.8	3.6	3750-0067		4.0	3.0
3332-1195		5.7	4.6	3771-7523		13	4.8
3334-4922		4.1	3.6	3909-9107		6.7	4.3
3334-4625		2.1	2.0	3909-9109		3.1	2.5
				3909-9113		2.3	2.6

Chem Div	Structure	Ratio to Control	
		1	2
3909-9114		3.4	3.5
3932-0079		9.9	7.0
3909-9134		5.5	5.3
3993-1905		2.4	2.2
3993-1906		11	10
3909-9877		2.3	2.2
4049-0294		4.7	2.5
4052-2657		7.4	4.8
4112-0771 CM001		4.9	5.1
4112-0824		6.6	6.5
4236-0892		2.3	2.0

Chem Div	Structure	Ratio to Control	
		1	2
4181-1899 CM008		2.1	2.3
4236-0953		3.9	14
4281-0310 CM022		23	16
4358-4339 CM016		2.2	2.0
4472-0052		2.7	6.6
4552-5135 CM032		2.0	3.5
4555-0196		2.1	2.1
4554-6912		4.5	2.6
4555-3401		3.0	2.4
4554-6941		3.4	2.8
4555-0120		2.9	2.3

Chem Div	Structure	Ratio to Control		Chem Div	Structure	Ratio to Control	
		1	2			1	2
4554-6856		3.1	2.3	5183-0418		2.3	2.5
4554-6933		2.9	2.3	5186-0442		5.5	2.2
4555-3107		8.2	4.9	5228-0312		12	7.0
4780-0024 CM033		3.8	2.9	5339-0215		4.8	3.5
4903-2108		2.0	2.4	5339-0193		5.8	2.4
4903-2155 CM002		3.1	2.8	5339-0551 CM051		5.3	5.1
4903-2156		3.1	3.2	5519-0751		3.8	2.2
5055-7414		3.1	2.6	5521-0170 CM034		2.7	2.1
5067-0687		8.3	6.3	5519-0629 CM010		2.6	2.3
5067-0952 CM014		4.2	4.5	5591-1893 CM047		8.2	13
5067-0859		12	12	5611-1897		2.3	2.5

Chem Div	Structure	Ratio to Control	
		1	2
5591-1925		8.6	3.1
5591-0588 CM006		2.6	2.5
5611-1894 CM004		5.4	5.2
5611-1978 CM005		4.3	3.6
5665-0118 CM019		3.6	3.6
5672-0068		2.3	2.0
5611-1984		3.0	2.6
5611-1987		5.3	5.7
5664-0379 CM035		2.5	4.0
5720-0005		2.5	2.4
5922-0026		7.6	6.2

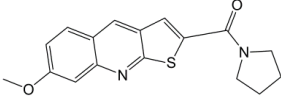
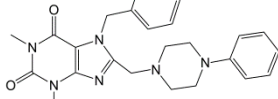
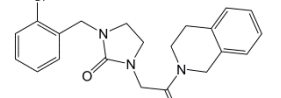
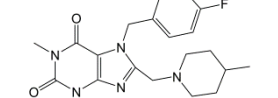
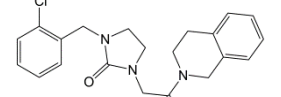
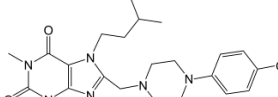
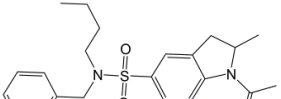
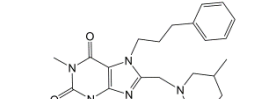
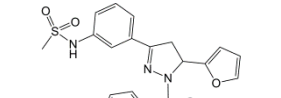
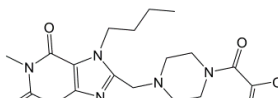
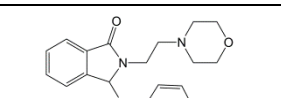
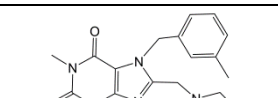
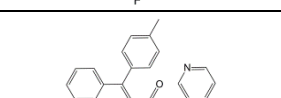
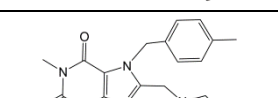
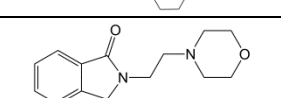
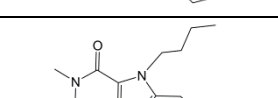
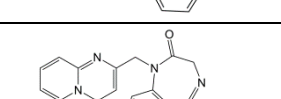
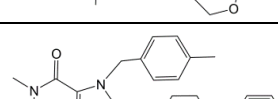
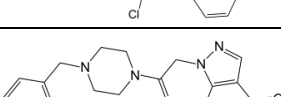
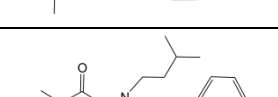
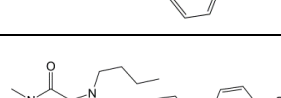
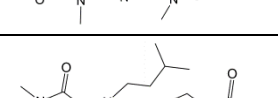
Chem Div	Structure	Ratio to Control	
		1	2
5922-0032		6.5	6.1
5922-0076 CM036		3.1	2.6
5941-1237 CM023		3.0	2.3
5946-0003		4.9	3.8
6049-1331		8.0	3.8
6049-1282 CM013		3.7	3.1
6049-0440		2.1	2.2
6049-2856		3.6	3.2
6049-2858		3.1	2.8
6113-1671		9.5	6.7
6113-2682		14	9.7
6228-1380 CM048		3.9	2.8



Chem Div	Structure	Ratio to Control	
		1	2
8009-2086		9.3	4.5
8012-8719		3.6	2.3
8013-5620		3.3	2.1
8013-5576		3.7	3.9
A001-5851		6.7	4.1
8015-3130		5.1	2.1
C066-0331		2.9	2.9
C066-0314		2.3	2.1
C077-0019		2.6	2.6
C085-2100		2.7	2.1
C109-0828		2.8	2.2

Chem Div	Structure	Ratio to Control	
		1	2
C109-0868 CM020		2.2	2.1
C123-0492 CM021		6.3	4.5
C156-0107 CM018		3.6	3.4
C174-0133 CM043		3.7	2.3
C174-0218 CM011		3.7	3.1
C174-0331 CM012		2.7	2.5
C174-0222		2.2	2.3
C187-0019		3.2	2.6
C200-0216		2.9	2.5
C200-1296		4.0	2.1
C269-0038		8.7	2.3



Chem Div	Structure	Ratio to Control		Chem Div	Structure	Ratio to Control	
		1	2			1	2
C303-0270		4.4	2.1	C470-0085		11	7.1
C303-0317		3.5	2.0	C470-0125		2.3	2.0
C305-0458		3.3	2.6	C470-0665		6.8	2.9
C361-0197 CM044		6.3	2.1	C470-0728		4.2	2.5
C366-0190 CM045		3.5	2.1	C470-0070		3.9	2.6
C429-0235		2.7	2.3	C470-0242		5.8	4.9
C378-0506		4.8	2.2	C470-0315		2.9	2.1
C429-0223		3.2	2.6	C470-0057		2.7	2.5
C386-0528		3.5	3.2	C470-0325 CM025		2.2	2.4
C434-0076 CM017		3.5	2.6	C470-0645 CM054		4.5	3.0
C470-0068 CM024		4.4	2.0	C470-0667 CM026		8.3	5.2

Chem Div	Structure	Ratio to Control	
		1	2
C470-0730		15	8.4
C470-0075		2.6	2.2
C470-0089		6.4	2.5
C470-0216		8.0	4.2
C470-0243		9.6	7.5
C470-0262 CM055		16	13
C470-0646 CM053		14	11
C470-0654 CM056		6.3	4.3
C470-0668		4.6	3.9
C470-0077		9.6	5.2
C470-0094		10	3.7

Chem Div	Structure	Ratio to Control	
		1	2
C470-0197		20	2.3
C470-0244		3.2	2.0
C470-0264		14	4.3
C470-0317		7.7	6.1
C470-0334		7.1	5.3
C470-0647		3.9	2.7
C470-0720		9.2	9.3
C470-0489 CM027		3.8	4.0
C470-0245		11	9.9
C470-0048		3.5	2.8
C470-0648		3.9	2.5

Chem Div	Structure	Ratio to Control		Chem Div	Structure	Ratio to Control	
		1	2			1	2
C470-0721 CM028		8.0	5.3	C470-0204		2.3	2.1
C470-0246 CM029		2.4	2.6	C470-0249		16	12
C470-0061		9.8	5.0	C470-0312		2.2	4.1
C470-0657 CM030		2.5	2.3	C470-0053		2.9	4.3
C470-0723		6.1	3.5	C470-0065 CM031		3.6	2.3
C470-0248		4.1	2.5	C470-0662		7.0	2.4
C470-0051		2.0	2.3	C470-0482		18	10
C470-0064		2.9	2.7	C470-0726		9.1	2.7
C470-0650 CM057		7.6	4.5	C593-0746		3.4	2.5
C470-0658		13	7.7	C598-0399		17	3.5
C470-0725		7.1	5.0	C598-0444		3.6	2.9

Chem Div	Structure	Ratio to Control		Chem Div	Structure	Ratio to Control	
		1	2			1	2
C629-0095		5.8	3.4	C720-0002		5.8	4.1
C629-0084		4.9	3.3	K263-0587 CM037		2.2	2.0
C629-0100 CM038		3.3	2.4	K284-1521		4.3	2.1
C660-0239		7.7	3.6	K284-1536		7.9	4.8
C692-0206		2.7	2.1	K284-7015		6.9	2.1
C692-0425		3.5	2.5	K297-0386		3.2	2.6
C692-0426		6.4	2.3	K409-0004		2.2	2.1
C692-0428		3.2	2.6	K643-0219		8.4	3.8
C692-0319		4.4	2.3	K781-5423 CM039		3.5	2.1
C692-0248		5.2	2.8	K781-5628		8.4	4.0
C692-0333		3.2	2.5	K783-4550		3.5	2.2

Chem Div	Structure	Ratio to Control	
		1	2
K784-6532		4.1	2.0
K784-5336		4.5	2.8
K784-9658		6.0	3.2
K784-8446		2.5	2.6
K784-9612		2.2	2.4
K784-9804		3.6	3.9
K785-0143 CM040		3.4	3.6
K785-0281		3.9	3.7
K785-0289		2.4	2.5
K784-9808		4.8	4.7
K785-0189		2.3	2.2

Chem Div	Structure	Ratio to Control	
		1	2
K786-9188 CM042		2.4	2.2
K786-9836 CM041		2.1	2.1
K788-0894		5.5	2.6
K788-6942		2.3	2.6
K788-1286		6.0	3.1
K788-8259		4.1	3.5
K788-7993		2.6	2.5
K894-1589		3.2	2.2
K913-0047		3.4	2.2
K913-0222		4.2	2.3
K913-0335		3.7	2.7
K913-0043		3.8	2.5

Supplementary Table 2: Inhibitors of ALDH1A1 esterase activity.

ChemDiv catalog number and structure of the 15 compounds that inhibited ALDH1A1 esterase activity in the HTS. The ratio to control value for the initial screen (1) and verification screen (2) are shown. The names of the compounds whose effect on aldehyde oxidation have been tested are shown as CM3XX below the Chem Div number. For three HTS compounds, close analogs of the compounds were used due to non-availability of hit compound.

Chem Div	Structure	Ratio to Control		Chem Div	Structure	Ratio to Control	
		1	2			1	2
3343-2924		0.4	0.5	C226-4110 CM303		0.4	0.4
2188-3271 CM310		Analog to 3343-2924		C226-4114		0.3	0.5
3998-0079		0.0	0.5	C521-0306		0.3	0.5
5311-0360 CM301		0.4	0.3	C548-3890 CM304		0.4	0.1
8008-9631 CM308		0.3	0.3	C638-0449 CM305		0.3	0.4
8008-0704 CM309		0.0	0.0	C699-0515		0.4	0.4
8011-8027 CM302		0.4	0.5	C699-0244 CM306		Analog to C699-0515	
8013-9940		0.3	0.5	K788-2754		0.5	0.5
C189-0365		0.3	0.5	K938-0803 CM307		Analog to K788-2754	

## References

1. O'Brien, P. J.; Siraki, A. G.; Shangari, N. Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health. *Crit Rev Toxicol* **2005**, *35*, 609-62.
2. Dockham, P. A.; Lee, M. O.; Sladek, N. E. Identification of human liver aldehyde dehydrogenases that catalyze the oxidation of aldophosphamide and retinaldehyde. *Biochem Pharmacol* **1992**, *43*, 2453-69.
3. Emadi, A.; Jones, R. J.; Brodsky, R. A. Cyclophosphamide and cancer: golden anniversary. *Nat Rev Clin Oncol* **2009**, *6*, 638-47.
4. Blackwell, L. F.; Bennett, A. F.; Buckley, P. D. Relationship between the mechanisms of the esterase and dehydrogenase activities of the cytoplasmic aldehyde dehydrogenase from sheep liver. An alternative view. *Biochemistry* **1983**, *22*, 3784-91.
5. Walsh, J. S.; Reese, M. J.; Thurmond, L. M. The metabolic activation of abacavir by human liver cytosol and expressed human alcohol dehydrogenase isozymes. *Chem Biol Interact* **2002**, *142*, 135-54.
6. Marchitti, S. A.; Brocker, C.; Stagos, D.; Vasiliou, V. Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. *Expert Opin Drug Metab Toxicol* **2008**, *4*, 697-720.
7. Berg, J. M.; Tymoczko, J. L.; Stryer, L.; Stryer, L. *Biochemistry*. 5th ed.; W.H. Freeman: New York, 2002.
8. Lodish, H. F. *Molecular cell biology*. 5th ed.; W.H. Freeman and Company: New York, 2003; p xxxiii, 973 , 79 p.
9. Curtis, J. M.; Hahn, W. S.; Long, E. K.; Burrill, J. S.; Arriaga, E. A.; Bernlohr, D. A. Protein carbonylation and metabolic control systems. *Trends Endocrinol Metab* **2012**, *23*, 399-406.
10. Ross, M.; Matthews, A.; Mangum, L. Chemical Atherogenesis: Role of Endogenous and Exogenous Poisons in Disease Development. *Toxics* **2014**, *2*, 17-34.
11. Dalle-Donne, I.; Aldini, G.; Carini, M.; Colombo, R.; Rossi, R.; Milzani, A. Protein carbonylation, cellular dysfunction, and disease progression. *J Cell Mol Med* **2006**, *10*, 389-406.
12. Brooks, P. J.; Zakhari, S. Acetaldehyde and the genome: beyond nuclear DNA adducts and carcinogenesis. *Environ Mol Mutagen* **2014**, *55*, 77-91.
13. Brooks, P. J.; Theruvathu, J. A. DNA adducts from acetaldehyde: implications for alcohol-related carcinogenesis. *Alcohol* **2005**, *35*, 187-93.

14. McMahon, P. B. *Chapter 5: Alcohols, aldehydes, and ketones*. Academic Press: New York, 1982; p xiii, 375 p.
15. Lin, S. J.; Guarente, L. Nicotinamide adenine dinucleotide, a metabolic regulator of transcription, longevity and disease. *Curr Opin Cell Biol* **2003**, 15, 241-6.
16. Jin, Y.; Penning, T. M. Aldo-keto reductases and bioactivation/detoxication. *Annu Rev Pharmacol Toxicol* **2007**, 47, 263-92.
17. Beedham, C.; Peet, C. F.; Panoutsopoulos, G. I.; Carter, H.; Smith, J. A. Role of aldehyde oxidase in biogenic amine metabolism. *Prog Brain Res* **1995**, 106, 345-53.
18. Vasiliou, V.; Pappa, A.; Petersen, D. R. Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chem Biol Interact* **2000**, 129, 1-19.
19. Berhane, K.; Widersten, M.; Engstrom, A.; Kozarich, J. W.; Mannervik, B. Detoxication of base propenals and other alpha, beta-unsaturated aldehyde products of radical reactions and lipid peroxidation by human glutathione transferases. *Proc Natl Acad Sci U S A* **1994**, 91, 1480-4.
20. Jackson, B.; Brocker, C.; Thompson, D. C.; Black, W.; Vasiliou, K.; Nebert, D. W.; Vasiliou, V. Update on the aldehyde dehydrogenase gene (ALDH) superfamily. *Hum Genomics* **2011**, 5, 283-303.
21. Vasiliou, V.; Bairoch, A.; Tipton, K. F.; Nebert, D. W. Eukaryotic aldehyde dehydrogenase (ALDH) genes: human polymorphisms, and recommended nomenclature based on divergent evolution and chromosomal mapping. *Pharmacogenetics* **1999**, 9, 421-34.
22. Vasiliou, V.; Thompson, D. C.; Smith, C.; Fujita, M.; Chen, Y. Aldehyde dehydrogenases: from eye crystallins to metabolic disease and cancer stem cells. *Chem Biol Interact* **2013**, 202, 2-10.
23. Koppaka, V.; Thompson, D. C.; Chen, Y.; Ellermann, M.; Nicolaou, K. C.; Juvonen, R. O.; Petersen, D.; Deitrich, R. A.; Hurley, T. D.; Vasiliou, V. Aldehyde dehydrogenase inhibitors: a comprehensive review of the pharmacology, mechanism of action, substrate specificity, and clinical application. *Pharmacol Rev* **2012**, 64, 520-39.
24. Feldman, R. I.; Weiner, H. Horse liver aldehyde dehydrogenase. II. Kinetics and mechanistic implications of the dehydrogenase and esterase activity. *J Biol Chem* **1972**, 247, 267-72.
25. Deak, K. L.; Dickerson, M. E.; Linney, E.; Enterline, D. S.; George, T. M.; Melvin, E. C.; Graham, F. L.; Siegel, D. G.; Hammock, P.; Mehlretter, L.; Bassuk, A. G.;



- Kessler, J. A.; Gilbert, J. R.; Speer, M. C. Analysis of ALDH1A2, CYP26A1, CYP26B1, CRABP1, and CRABP2 in human neural tube defects suggests a possible association with alleles in ALDH1A2. *Birth Defects Res A Clin Mol Teratol* **2005**, 73, 868-75.
26. Estey, T.; Cantore, M.; Weston, P. A.; Carpenter, J. F.; Petrash, J. M.; Vasiliou, V. Mechanisms involved in the protection of UV-induced protein inactivation by the corneal crystallin ALDH3A1. *J Biol Chem* **2007**, 282, 4382-92.
27. Chen, Y.; Mehta, G.; Vasiliou, V. Antioxidant defenses in the ocular surface. *Ocul Surf* **2009**, 7, 176-85.
28. Lassen, N.; Bateman, J. B.; Estey, T.; Kuszak, J. R.; Nees, D. W.; Piatigorsky, J.; Duester, G.; Day, B. J.; Huang, J.; Hines, L. M.; Vasiliou, V. Multiple and additive functions of ALDH3A1 and ALDH1A1: cataract phenotype and ocular oxidative damage in *Aldh3a1(-/-)/Aldh1a1(-/-)* knock-out mice. *J Biol Chem* **2007**, 282, 25668-76.
29. Vasiliou, V.; Pappa, A.; Estey, T. Role of human aldehyde dehydrogenases in endobiotic and xenobiotic metabolism. *Drug Metab Rev* **2004**, 36, 279-99.
30. Burke, W. J.; Li, S. W.; Williams, E. A.; Nonneman, R.; Zahm, D. S. 3,4-Dihydroxyphenylacetaldehyde is the toxic dopamine metabolite in vivo: implications for Parkinson's disease pathogenesis. *Brain Res* **2003**, 989, 205-13.
31. Yao, L.; Fan, P.; Arolfo, M.; Jiang, Z.; Olive, M. F.; Zablocki, J.; Sun, H. L.; Chu, N.; Lee, J.; Kim, H. Y.; Leung, K.; Shryock, J.; Blackburn, B.; Diamond, I. Inhibition of aldehyde dehydrogenase-2 suppresses cocaine seeking by generating THP, a cocaine use-dependent inhibitor of dopamine synthesis. *Nat Med* **2010**, 16, 1024-8.
32. Akaboshi, S.; Hogema, B. M.; Novelletto, A.; Malaspina, P.; Salomons, G. S.; Maropoulos, G. D.; Jakobs, C.; Grompe, M.; Gibson, K. M. Mutational spectrum of the succinate semialdehyde dehydrogenase (ALDH5A1) gene and functional analysis of 27 novel disease-causing mutations in patients with SSADH deficiency. *Hum Mutat* **2003**, 22, 442-50.
33. Rizzo, W. B.; Carney, G. Sjogren-Larsson syndrome: diversity of mutations and polymorphisms in the fatty aldehyde dehydrogenase gene (ALDH3A2). *Hum Mutat* **2005**, 26, 1-10.
34. Geraghty, M. T.; Vaughn, D.; Nicholson, A. J.; Lin, W. W.; Jimenez-Sanchez, G.; Obie, C.; Flynn, M. P.; Valle, D.; Hu, C. A. Mutations in the Delta1-pyrroline 5-carboxylate dehydrogenase gene cause type II hyperprolinemia. *Hum Mol Genet* **1998**, 7, 1411-5.

35. Kamoun, P.; Aral, B.; Saudubray, J. M. [A new inherited metabolic disease: delta1-pyrroline 5-carboxylate synthetase deficiency]. *Bull Acad Natl Med* **1998**, 182, 131-7; discussion 138-9.
36. Vasiliou, V.; Sandoval, M.; Backos, D. S.; Jackson, B. C.; Chen, Y.; Reigan, P.; Lanaspá, M. A.; Johnson, R. J.; Koppaka, V.; Thompson, D. C. ALDH16A1 is a novel non-catalytic enzyme that may be involved in the etiology of gout via protein-protein interactions with HPRT1. *Chem Biol Interact* **2013**, 202, 22-31.
37. Perozich, J.; Nicholas, H.; Wang, B. C.; Lindahl, R.; Hempel, J. Relationships within the aldehyde dehydrogenase extended family. *Protein Sci* **1999**, 8, 137-46.
38. Farres, J.; Wang, T. T.; Cunningham, S. J.; Weiner, H. Investigation of the active site cysteine residue of rat liver mitochondrial aldehyde dehydrogenase by site-directed mutagenesis. *Biochemistry* **1995**, 34, 2592-8.
39. Wang, X.; Weiner, H. Involvement of glutamate 268 in the active site of human liver mitochondrial (class 2) aldehyde dehydrogenase as probed by site-directed mutagenesis. *Biochemistry* **1995**, 34, 237-43.
40. Blackwell, L. F.; Motion, R. L.; MacGibbon, A. K.; Hardman, M. J.; Buckley, P. D. Evidence that the slow conformation change controlling NADH release from the enzyme is rate-limiting during the oxidation of propionaldehyde by aldehyde dehydrogenase. *Biochem J* **1987**, 242, 803-8.
41. Weiner, H.; Hu, J. H.; Sanny, C. G. Rate-limiting steps for the esterase and dehydrogenase reaction catalyzed by horse liver aldehyde dehydrogenase. *J Biol Chem* **1976**, 251, 3853-5.
42. Mann, C. J.; Weiner, H. Differences in the roles of conserved glutamic acid residues in the active site of human class 3 and class 2 aldehyde dehydrogenases. *Protein Sci* **1999**, 8, 1922-9.
43. Sidhu, R. S.; Blair, A. H. Human liver aldehyde dehydrogenase. Esterase activity. *J Biol Chem* **1975**, 250, 7894-8.
44. Liu, Z. J.; Sun, Y. J.; Rose, J.; Chung, Y. J.; Hsiao, C. D.; Chang, W. R.; Kuo, I.; Perozich, J.; Lindahl, R.; Hempel, J.; Wang, B. C. The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD and the Rossmann fold. *Nat Struct Biol* **1997**, 4, 317-26.
45. Steinmetz, C. G.; Xie, P.; Weiner, H.; Hurley, T. D. Structure of mitochondrial aldehyde dehydrogenase: the genetic component of ethanol aversion. *Structure* **1997**, 5, 701-11.

46. Ni, L.; Zhou, J.; Hurley, T. D.; Weiner, H. Human liver mitochondrial aldehyde dehydrogenase: three-dimensional structure and the restoration of solubility and activity of chimeric forms. *Protein Sci* **1999**, *8*, 2784-90.
47. Wang, M. F.; Han, C. L.; Yin, S. J. Substrate specificity of human and yeast aldehyde dehydrogenases. *Chem Biol Interact* **2009**, *178*, 36-9.
48. Chambon, P. A decade of molecular biology of retinoic acid receptors. *FASEB J* **1996**, *10*, 940-54.
49. King, G.; Holmes, R. Human corneal and lens aldehyde dehydrogenases. Purification and properties of human lens ALDH1 and differential expression as major soluble proteins in human lens (ALDH1) and cornea (ALDH3). *Adv Exp Med Biol* **1997**, *414*, 19-27.
50. Lassen, N.; Black, W. J.; Estey, T.; Vasiliou, V. The role of corneal crystallins in the cellular defense mechanisms against oxidative stress. *Semin Cell Dev Biol* **2008**, *19*, 100-12.
51. Zhou, J.; Weiner, H. Binding of thyroxine analogs to human liver aldehyde dehydrogenases. *Eur J Biochem* **1997**, *245*, 123-8.
52. Deng, L.; Shipley, G. L.; Loose-Mitchell, D. S.; Stancel, G. M.; Broaddus, R.; Pickar, J. H.; Davies, P. J. Coordinate regulation of the production and signaling of retinoic acid by estrogen in the human endometrium. *J Clin Endocrinol Metab* **2003**, *88*, 2157-63.
53. Kumar, S.; Sandell, L. L.; Trainor, P. A.; Koentgen, F.; Duester, G. Alcohol and aldehyde dehydrogenases: Retinoid metabolic effects in mouse knockout models. *Biochim Biophys Acta* **2012**, *1821*, 198-205.
54. Niederreither, K.; Fraulob, V.; Garnier, J. M.; Chambon, P.; Dolle, P. Differential expression of retinoic acid-synthesizing (RALDH) enzymes during fetal development and organ differentiation in the mouse. *Mech Dev* **2002**, *110*, 165-71.
55. Theodosiou, M.; Laudet, V.; Schubert, M. From carrot to clinic: an overview of the retinoic acid signaling pathway. *Cell Mol Life Sci* **2010**, *67*, 1423-45.
56. Bhat, P. V.; Samaha, H. Kinetic properties of the human liver cytosolic aldehyde dehydrogenase for retinal isomers. *Biochem Pharmacol* **1999**, *57*, 195-7.
57. Lin, M.; Napoli, J. L. cDNA cloning and expression of a human aldehyde dehydrogenase (ALDH) active with 9-cis-retinal and identification of a rat ortholog, ALDH12. *J Biol Chem* **2000**, *275*, 40106-12.

58. Marcato, P.; Dean, C. A.; Giacomantonio, C. A.; Lee, P. W. Aldehyde dehydrogenase: its role as a cancer stem cell marker comes down to the specific isoform. *Cell Cycle* **2011**, *10*, 1378-84.
59. Ross, S. A.; McCaffery, P. J.; Drager, U. C.; De Luca, L. M. Retinoids in embryonal development. *Physiol Rev* **2000**, *80*, 1021-54.
60. AceView. <http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/>
61. Spence, J. P.; Liang, T.; Eriksson, C. J.; Taylor, R. E.; Wall, T. L.; Ehlers, C. L.; Carr, L. G. Evaluation of aldehyde dehydrogenase 1 promoter polymorphisms identified in human populations. *Alcohol Clin Exp Res* **2003**, *27*, 1389-94.
62. Lind, P. A.; Eriksson, C. J.; Wilhelmsen, K. C. The role of aldehyde dehydrogenase-1 (ALDH1A1) polymorphisms in harmful alcohol consumption in a Finnish population. *Hum Genomics* **2008**, *3*, 24-35.
63. Schafer, A.; Teufel, J.; Ringel, F.; Bettstetter, M.; Hoepner, I.; Rasper, M.; Gempt, J.; Koeritzer, J.; Schmidt-Graf, F.; Meyer, B.; Beier, C. P.; Schlegel, J. Aldehyde dehydrogenase 1A1--a new mediator of resistance to temozolomide in glioblastoma. *Neuro Oncol* **2012**, *14*, 1452-64.
64. Michotte, A. Recent developments in the neuropathological diagnosis of Parkinson's disease and parkinsonism. *Acta Neurol Belg* **2003**, *103*, 155-8.
65. Marsden, C. D. Parkinson's disease. *Lancet* **1990**, *335*, 948-52.
66. Bernheimer, H.; Birkmayer, W.; Hornykiewicz, O.; Jellinger, K.; Seitelberger, F. Brain dopamine and the syndromes of Parkinson and Huntington. Clinical, morphological and neurochemical correlations. *J Neurol Sci* **1973**, *20*, 415-55.
67. Marchitti, S. A.; Deitrich, R. A.; Vasiliou, V. Neurotoxicity and metabolism of the catecholamine-derived 3,4-dihydroxyphenylacetaldehyde and 3,4-dihydroxyphenylglycolaldehyde: the role of aldehyde dehydrogenase. *Pharmacol Rev* **2007**, *59*, 125-50.
68. Tarazi, F. I.; Sahli, Z. T.; Wolny, M.; Mousa, S. A. Emerging therapies for Parkinson's disease: From bench to bedside. *Pharmacol Ther* **2014**, *144*, 123-133.
69. Jankovic, J. Motor fluctuations and dyskinesias in Parkinson's disease: clinical manifestations. *Mov Disord* **2005**, *20* Suppl 11, S11-6.
70. Mandel, S.; Grunblatt, E.; Riederer, P.; Amariglio, N.; Jacob-Hirsch, J.; Rechavi, G.; Youdim, M. B. Gene expression profiling of sporadic Parkinson's disease substantia nigra

pars compacta reveals impairment of ubiquitin-proteasome subunits, SKP1A, aldehyde dehydrogenase, and chaperone HSC-70. *Ann N Y Acad Sci* **2005**, 1053, 356-75.

71. Galter, D.; Buervenich, S.; Carmine, A.; Anvret, M.; Olson, L. ALDH1 mRNA: presence in human dopamine neurons and decreases in substantia nigra in Parkinson's disease and in the ventral tegmental area in schizophrenia. *Neurobiol Dis* **2003**, 14, 637-47.

72. Werner, C. J.; Heyny-von Haussen, R.; Mall, G.; Wolf, S. Proteome analysis of human substantia nigra in Parkinson's disease. *Proteome Sci* **2008**, 6, 8.

73. MacKerell, A. D., Jr.; Pietruszko, R. Chemical modification of human aldehyde dehydrogenase by physiological substrate. *Biochim Biophys Acta* **1987**, 911, 306-17.

74. Wey, M. C.; Fernandez, E.; Martinez, P. A.; Sullivan, P.; Goldstein, D. S.; Strong, R. Neurodegeneration and motor dysfunction in mice lacking cytosolic and mitochondrial aldehyde dehydrogenases: implications for Parkinson's disease. *PLoS One* **2012**, 7, e31522.

75. Chen, C. H.; Budas, G. R.; Churchill, E. N.; Disatnik, M. H.; Hurley, T. D.; Mochly-Rosen, D. Activation of aldehyde dehydrogenase-2 reduces ischemic damage to the heart. *Science* **2008**, 321, 1493-5.

76. Kastan, M. B.; Schlaffer, E.; Russo, J. E.; Colvin, O. M.; Civin, C. I.; Hilton, J. Direct demonstration of elevated aldehyde dehydrogenase in human hematopoietic progenitor cells. *Blood* **1990**, 75, 1947-50.

77. Jones, R. J.; Barber, J. P.; Vala, M. S.; Collector, M. I.; Kaufmann, S. H.; Ludeman, S. M.; Colvin, O. M.; Hilton, J. Assessment of aldehyde dehydrogenase in viable cells. *Blood* **1995**, 85, 2742-6.

78. Lapidot, T.; Sirard, C.; Vormoor, J.; Murdoch, B.; Hoang, T.; Caceres-Cortes, J.; Minden, M.; Paterson, B.; Caligiuri, M. A.; Dick, J. E. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **1994**, 367, 645-8.

79. Al-Hajj, M.; Wicha, M. S.; Benito-Hernandez, A.; Morrison, S. J.; Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* **2003**, 100, 3983-8.

80. Wicha, M. S.; Liu, S.; Dontu, G. Cancer stem cells: an old idea--a paradigm shift. *Cancer Res* **2006**, 66, 1883-90; discussion 1895-6.

81. Reya, T.; Morrison, S. J.; Clarke, M. F.; Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **2001**, 414, 105-11.

82. Levi, B. P.; Yilmaz, O. H.; Duester, G.; Morrison, S. J. Aldehyde dehydrogenase 1a1 is dispensable for stem cell function in the mouse hematopoietic and nervous systems. *Blood* **2009**, 113, 1670-80.
83. Hu, X. T.; Zuckerman, K. S. Role of cell cycle regulatory molecules in retinoic acid- and vitamin D3-induced differentiation of acute myeloid leukaemia cells. *Cell Prolif* **2014**, 47, 200-10.
84. Muramoto, G. G.; Russell, J. L.; Safi, R.; Salter, A. B.; Himburg, H. A.; Daher, P.; Meadows, S. K.; Doan, P.; Storms, R. W.; Chao, N. J.; McDonnell, D. P.; Chute, J. P. Inhibition of aldehyde dehydrogenase expands hematopoietic stem cells with radioprotective capacity. *Stem Cells* **2010**, 28, 523-34.
85. Chute, J. P.; Muramoto, G. G.; Whitesides, J.; Colvin, M.; Safi, R.; Chao, N. J.; McDonnell, D. P. Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells. *Proc Natl Acad Sci U S A* **2006**, 103, 11707-12.
86. Hayashi, K.; Yokozaki, H.; Naka, K.; Yasui, W.; Lotan, R.; Tahara, E. Overexpression of retinoic acid receptor beta induces growth arrest and apoptosis in oral cancer cell lines. *Jpn J Cancer Res* **2001**, 92, 42-50.
87. Lokshin, A.; Zhang, H.; Mayotte, J.; Lokshin, M.; Levitt, M. L. Early effects of retinoic acid on proliferation, differentiation and apoptosis in non-small cell lung cancer cell lines. *Anticancer Res* **1999**, 19, 5251-4.
88. Kim, D. G.; Jo, B. H.; You, K. R.; Ahn, D. S. Apoptosis induced by retinoic acid in Hep 3B cells in vitro. *Cancer Lett* **1996**, 107, 149-59.
89. Polakis, P. Wnt signaling and cancer. *Genes Dev* **2000**, 14, 1837-51.
90. Weinbach, E. C.; Garbus, J. Mechanism of action of reagents that uncouple oxidative phosphorylation. *Nature* **1969**, 221, 1016-8.
91. Wieland, A.; Trageser, D.; Gogolok, S.; Reinartz, R.; Hofer, H.; Keller, M.; Leinhaas, A.; Schelle, R.; Normann, S.; Klaas, L.; Waha, A.; Koch, P.; Fimmers, R.; Pietsch, T.; Yachnis, A. T.; Pincus, D. W.; Steindler, D. A.; Brustle, O.; Simon, M.; Glas, M.; Scheffler, B. Anticancer effects of niclosamide in human glioblastoma. *Clin Cancer Res* **2013**, 19, 4124-36.
92. Yo, Y. T.; Lin, Y. W.; Wang, Y. C.; Balch, C.; Huang, R. L.; Chan, M. W.; Sytwu, H. K.; Chen, C. K.; Chang, C. C.; Nephew, K. P.; Huang, T.; Yu, M. H.; Lai, H. C. Growth inhibition of ovarian tumor-initiating cells by niclosamide. *Mol Cancer Ther* **2012**, 11, 1703-12.

93. Dontu, G.; Abdallah, W. M.; Foley, J. M.; Jackson, K. W.; Clarke, M. F.; Kawamura, M. J.; Wicha, M. S. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev* **2003**, *17*, 1253-70.
94. Saw, Y. T.; Yang, J.; Ng, S. K.; Liu, S.; Singh, S.; Singh, M.; Welch, W. R.; Tsuda, H.; Fong, W. P.; Thompson, D.; Vasiliou, V.; Berkowitz, R. S.; Ng, S. W. Characterization of aldehyde dehydrogenase isozymes in ovarian cancer tissues and sphere cultures. *BMC Cancer* **2012**, *12*, 329.
95. Charafe-Jauffret, E.; Ginestier, C.; Iovino, F.; Tarpin, C.; Diebel, M.; Esterni, B.; Houvenaeghel, G.; Extra, J. M.; Bertucci, F.; Jacquemier, J.; Xerri, L.; Dontu, G.; Stassi, G.; Xiao, Y.; Barsky, S. H.; Birnbaum, D.; Viens, P.; Wicha, M. S. Aldehyde dehydrogenase 1-positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer. *Clin Cancer Res* **2010**, *16*, 45-55.
96. Bortolomai, I.; Canevari, S.; Facetti, I.; De Cecco, L.; Castellano, G.; Zacchetti, A.; Alison, M. R.; Miotti, S. Tumor initiating cells: development and critical characterization of a model derived from the A431 carcinoma cell line forming spheres in suspension. *Cell Cycle* **2010**, *9*, 1194-206.
97. Lukacs, R. U.; Lawson, D. A.; Xin, L.; Zong, Y.; Garraway, I.; Goldstein, A. S.; Memarzadeh, S.; Witte, O. N. Epithelial stem cells of the prostate and their role in cancer progression. *Cold Spring Harb Symp Quant Biol* **2008**, *73*, 491-502.
98. Sodek, K. L.; Ringuette, M. J.; Brown, T. J. Compact spheroid formation by ovarian cancer cells is associated with contractile behavior and an invasive phenotype. *Int J Cancer* **2009**, *124*, 2060-70.
99. National Cancer Institute (U.S.). Cancer Statistics Branch. SEER cancer statistics review. In National Cancer Institute, Cancer Statistics Branch.: Bethesda, Md.
100. Group, U. S. C. S. W. United States Cancer Statistics: 1999–2011 Incidence and Mortality Web-based Report. . In Department of Health and Human Services, C. f. D. C. a. P., and National Cancer Institute, Ed. Atlanta, GA, 2014.
101. Condello, S.; Morgan, C. A.; Nagdas, S.; Cao, L.; Turek, J.; Hurley, T. D.; Matei, D. beta-Catenin-regulated ALDH1A1 is a target in ovarian cancer spheroids. *Oncogene* **2014**.
102. Wulf, G. G.; Wang, R. Y.; Kuehnle, I.; Weidner, D.; Marini, F.; Brenner, M. K.; Andreeff, M.; Goodell, M. A. A leukemic stem cell with intrinsic drug efflux capacity in acute myeloid leukemia. *Blood* **2001**, *98*, 1166-73.

103. Liu, G.; Yuan, X.; Zeng, Z.; Tunici, P.; Ng, H.; Abdulkadir, I. R.; Lu, L.; Irvin, D.; Black, K. L.; Yu, J. S. Analysis of gene expression and chemoresistance of CD133+ cancer stem cells in glioblastoma. *Mol Cancer* **2006**, *5*, 67.
104. Dalerba, P.; Cho, R. W.; Clarke, M. F. Cancer stem cells: models and concepts. *Annu Rev Med* **2007**, *58*, 267-84.
105. Hilton, J. Role of aldehyde dehydrogenase in cyclophosphamide-resistant L1210 leukemia. *Cancer Res* **1984**, *44*, 5156-60.
106. Parajuli, B.; Fishel, M. L.; Hurley, T. D. Selective ALDH3A1 inhibition by benzimidazole analogues increase mafosfamide sensitivity in cancer cells. *J Med Chem* **2014**, *57*, 449-61.
107. Parajuli, B.; Georgiadis, T. M.; Fishel, M. L.; Hurley, T. D. Development of selective inhibitors for human aldehyde dehydrogenase 3A1 (ALDH3A1) for the enhancement of cyclophosphamide cytotoxicity. *Chembiochem* **2014**, *15*, 701-12.
108. Arend, R. C.; Londono-Joshi, A. I.; Samant, R. S.; Li, Y.; Conner, M.; Hidalgo, B.; Alvarez, R. D.; Landen, C. N.; Straughn, J. M.; Buchsbaum, D. J. Inhibition of Wnt/beta-catenin pathway by niclosamide: a therapeutic target for ovarian cancer. *Gynecol Oncol* **2014**, *134*, 112-20.
109. Galic, S.; Oakhill, J. S.; Steinberg, G. R. Adipose tissue as an endocrine organ. *Mol Cell Endocrinol* **2010**, *316*, 129-39.
110. Saely, C. H.; Geiger, K.; Drexel, H. Brown versus white adipose tissue: a mini-review. *Gerontology* **2012**, *58*, 15-23.
111. Virtanen, K. A.; Lidell, M. E.; Orava, J.; Heglind, M.; Westergren, R.; Niemi, T.; Taittonen, M.; Laine, J.; Savisto, N. J.; Enerback, S.; Nuutila, P. Functional brown adipose tissue in healthy adults. *N Engl J Med* **2009**, *360*, 1518-25.
112. Cypess, A. M.; Lehman, S.; Williams, G.; Tal, I.; Rodman, D.; Goldfine, A. B.; Kuo, F. C.; Palmer, E. L.; Tseng, Y. H.; Doria, A.; Kolodny, G. M.; Kahn, C. R. Identification and importance of brown adipose tissue in adult humans. *N Engl J Med* **2009**, *360*, 1509-17.
113. van Marken Lichtenbelt, W. D.; Vanhomerig, J. W.; Smulders, N. M.; Drossaerts, J. M.; Kemerink, G. J.; Bouvy, N. D.; Schrauwen, P.; Teule, G. J. Cold-activated brown adipose tissue in healthy men. *N Engl J Med* **2009**, *360*, 1500-8.
114. Cristancho, A. G.; Lazar, M. A. Forming functional fat: a growing understanding of adipocyte differentiation. *Nat Rev Mol Cell Biol* **2011**, *12*, 722-34.



115. Klyde, B. J.; Hirsch, J. Increased cellular proliferation in adipose tissue of adult rats fed a high-fat diet. *J Lipid Res* **1979**, *20*, 705-15.
116. Faust, I. M.; Johnson, P. R.; Stern, J. S.; Hirsch, J. Diet-induced adipocyte number increase in adult rats: a new model of obesity. *Am J Physiol* **1978**, *235*, E279-86.
117. Johnson, P. R.; Hirsch, J. Cellularity of adipose depots in six strains of genetically obese mice. *J Lipid Res* **1972**, *13*, 2-11.
118. Ziouzenkova, O.; Orasanu, G.; Sharlach, M.; Akiyama, T. E.; Berger, J. P.; Viereck, J.; Hamilton, J. A.; Tang, G.; Dolnikowski, G. G.; Vogel, S.; Duester, G.; Plutzky, J. Retinaldehyde represses adipogenesis and diet-induced obesity. *Nat Med* **2007**, *13*, 695-702.
119. Kiefer, F. W.; Vernochet, C.; O'Brien, P.; Spoerl, S.; Brown, J. D.; Nallamshetty, S.; Zeyda, M.; Stulnig, T. M.; Cohen, D. E.; Kahn, C. R.; Plutzky, J. Retinaldehyde dehydrogenase 1 regulates a thermogenic program in white adipose tissue. *Nat Med* **2012**.
120. Berry, D. C.; Noy, N. All-trans-retinoic acid represses obesity and insulin resistance by activating both peroxisome proliferation-activated receptor beta/delta and retinoic acid receptor. *Mol Cell Biol* **2009**, *29*, 3286-96.
121. Reichert, B.; Yasmeen, R.; Jeyakumar, S. M.; Yang, F.; Thomou, T.; Alder, H.; Duester, G.; Maiseyeu, A.; Mihai, G.; Harrison, E. H.; Rajagopalan, S.; Kirkland, J. L.; Ziouzenkova, O. Concerted action of aldehyde dehydrogenases influences depot-specific fat formation. *Mol Endocrinol* **2011**, *25*, 799-809.
122. Storms, R. W.; Trujillo, A. P.; Springer, J. B.; Shah, L.; Colvin, O. M.; Ludeman, S. M.; Smith, C. Isolation of primitive human hematopoietic progenitors on the basis of aldehyde dehydrogenase activity. *Proc Natl Acad Sci U S A* **1999**, *96*, 9118-23.
123. Ginestier, C.; Hur, M. H.; Charafe-Jauffret, E.; Monville, F.; Dutcher, J.; Brown, M.; Jacquemier, J.; Viens, P.; Klier, C. G.; Liu, S.; Schott, A.; Hayes, D.; Birnbaum, D.; Wicha, M. S.; Dontu, G. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* **2007**, *1*, 555-67.
124. Russo, J. E.; Haugwitz, D.; Hilton, J. Inhibition of mouse cytosolic aldehyde dehydrogenase by 4-(diethylamino)benzaldehyde. *Biochem Pharmacol* **1988**, *37*, 1639-42.
125. Marcato, P.; Dean, C. A.; Pan, D.; Araslanova, R.; Gillis, M.; Joshi, M.; Helyer, L.; Pan, L.; Leidal, A.; Gujar, S.; Giacomantonio, C. A.; Lee, P. W. Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. *Stem Cells* **2011**, *29*, 32-45.

126. Moreb, J. S.; Ucar, D.; Han, S.; Amory, J. K.; Goldstein, A. S.; Ostmark, B.; Chang, L. J. The enzymatic activity of human aldehyde dehydrogenases 1A2 and 2 (ALDH1A2 and ALDH2) is detected by Aldefluor, inhibited by diethylaminobenzaldehyde and has significant effects on cell proliferation and drug resistance. *Chem Biol Interact* **2012**, 195, 52-60.
127. Morgan, C. A.; Parajuli, B.; Buchman, C.; Dria, K.; Hurley, T. D. N,N-diethylaminobenzaldehyde (DEAB) as a substrate and mechanism-based inhibitor for human ALDH isoenzymes. *Chemico-Biological Interactions* **2014**, in press.
128. Goedde, H. W.; Agarwal, D. P. Polymorphism of aldehyde dehydrogenase and alcohol sensitivity. *Enzyme* **1987**, 37, 29-44.
129. Conticello, C.; Martinetti, D.; Adamo, L.; Buccheri, S.; Giuffrida, R.; Parrinello, N.; Lombardo, L.; Anastasi, G.; Amato, G.; Cavalli, M.; Chiarenza, A.; De Maria, R.; Giustolisi, R.; Gulisano, M.; Di Raimondo, F. Disulfiram, an old drug with new potential therapeutic uses for human hematological malignancies. *Int J Cancer* **2012**, 131, 2197-203.
130. Liu, P.; Kumar, I. S.; Brown, S.; Kannappan, V.; Tawari, P. E.; Tang, J. Z.; Jiang, W.; Armesilla, A. L.; Darling, J. L.; Wang, W. Disulfiram targets cancer stem-like cells and reverses resistance and cross-resistance in acquired paclitaxel-resistant triple-negative breast cancer cells. *Br J Cancer* **2013**, 109, 1876-85.
131. Raha, D.; Wilson, T. R.; Peng, J.; Peterson, D.; Yue, P.; Evangelista, M.; Wilson, C.; Merchant, M.; Settleman, J. The cancer stem cell marker aldehyde dehydrogenase is required to maintain a drug-tolerant tumor cell subpopulation. *Cancer Res* **2014**, 74, 3579-90.
132. Ress, N. B.; Hailey, J. R.; Maronpot, R. R.; Bucher, J. R.; Travlos, G. S.; Haseman, J. K.; Orzech, D. P.; Johnson, J. D.; Hejtmancik, M. R. Toxicology and carcinogenesis studies of microencapsulated citral in rats and mice. *Toxicol Sci* **2003**, 71, 198-206.
133. Modak, T.; Mukhopadhaya, A. Effects of citral, a naturally occurring antiadipogenic molecule, on an energy-intense diet model of obesity. *Indian J Pharmacol* **2011**, 43, 300-5.
134. Staub, R. E.; Quistad, G. B.; Casida, J. E. Mechanism for benomyl action as a mitochondrial aldehyde dehydrogenase inhibitor in mice. *Chem Res Toxicol* **1998**, 11, 535-43.
135. Messiha, F. S. Effect of gossypol on kinetics of mouse liver alcohol and aldehyde dehydrogenase. *Gen Pharmacol* **1991**, 22, 573-6.

136. Parajuli, B.; Kimble-Hill, A. C.; Khanna, M.; Ivanova, Y.; Meroueh, S.; Hurley, T. D. Discovery of novel regulators of aldehyde dehydrogenase isoenzymes. *Chem Biol Interact* **2011**, 191, 153-8.
137. Khanna, M.; Chen, C. H.; Kimble-Hill, A.; Parajuli, B.; Perez-Miller, S.; Baskaran, S.; Kim, J.; Dria, K.; Vasiliou, V.; Mochly-Rosen, D.; Hurley, T. D. Discovery of a novel class of covalent inhibitor for aldehyde dehydrogenases. *J Biol Chem* **2011**, 286, 43486-94.
138. Luo, Y.; Dallaglio, K.; Chen, Y.; Robinson, W. A.; Robinson, S. E.; McCarter, M. D.; Wang, J.; Gonzalez, R.; Thompson, D. C.; Norris, D. A.; Roop, D. R.; Vasiliou, V.; Fujita, M. ALDH1A isozymes are markers of human melanoma stem cells and potential therapeutic targets. *Stem Cells* **2012**, 30, 2100-13.
139. Zorn, J. A.; Wells, J. A. Turning enzymes ON with small molecules. *Nat Chem Biol* **2010**, 6, 179-188.
140. Perez-Miller, S.; Younus, H.; Vanam, R.; Chen, C. H.; Mochly-Rosen, D.; Hurley, T. D. Alda-1 is an agonist and chemical chaperone for the common human aldehyde dehydrogenase 2 variant. *Nat Struct Mol Biol* **2010**, 17, 159-64.
141. Kotraiah, V.; Pallares, D.; Toema, D.; Kong, D.; Beausoleil, E. Identification of aldehyde dehydrogenase 1A1 modulators using virtual screening. *J Enzyme Inhib Med Chem* **2013**, 28, 489-94.
142. Fitzmaurice, A. G.; Rhodes, S. L.; Lulla, A.; Murphy, N. P.; Lam, H. A.; O'Donnell, K. C.; Barnhill, L.; Casida, J. E.; Cockburn, M.; Sagasti, A.; Stahl, M. C.; Maidment, N. T.; Ritz, B.; Bronstein, J. M. Aldehyde dehydrogenase inhibition as a pathogenic mechanism in Parkinson disease. *Proc Natl Acad Sci U S A* **2013**, 110, 636-41.
143. Dixon, M.; Webb, E. C. *Enzymes*. 3d ed.; Academic Press: New York, 1979; p xxiv, 1116 p.
144. Rossmann, M. G.; Liljas, A.; Branden, C.-I.; Banaszak, L. J. Evolutionary and Structural Relationships among Dehydrogenases. In *The Enzymes*, Boyer, P. D., Ed. Academic Press: New York City, 1975; Vol. 11, pp 61-102.
145. Gonzalez-Segura, L.; Rudino-Pinera, E.; Munoz-Clares, R. A.; Horjales, E. The crystal structure of a ternary complex of betaine aldehyde dehydrogenase from *Pseudomonas aeruginosa* Provides new insight into the reaction mechanism and shows a novel binding mode of the 2'-phosphate of NADP<sup>+</sup> and a novel cation binding site. *J Mol Biol* **2009**, 385, 542-57.
146. Tsybovsky, Y.; Donato, H.; Krupenko, N. I.; Davies, C.; Krupenko, S. A. Crystal structures of the carboxyl terminal domain of rat 10-formyltetrahydrofolate

dehydrogenase: implications for the catalytic mechanism of aldehyde dehydrogenases. *Biochemistry* **2007**, 46, 2917-29.

147. Keung, W. M.; Vallee, B. L. Daidzin: a potent, selective inhibitor of human mitochondrial aldehyde dehydrogenase. *Proc Natl Acad Sci U S A* **1993**, 90, 1247-51.

148. Baker, H. M.; Brown, R. L.; Dobbs, A. J.; Blackwell, L. F.; Buckley, P. D.; Hardman, M. J.; Hill, J. P.; Kitson, K. E.; Kitson, T. M.; Baker, E. N. Crystallization and preliminary X-ray diffraction studies on cytosolic (class 1) aldehyde dehydrogenase from sheep liver. *J Mol Biol* **1994**, 241, 263-4.

149. Baker, H. M.; Brown, R. L.; Dobbs, A. J.; Kitson, K. E.; Kitson, T. M.; Baker, E. N. Crystallization of sheep liver cytosolic aldehyde dehydrogenase in a form suitable for high resolution X-ray structural analysis. *Adv Exp Med Biol* **1995**, 372, 67-70.

150. Perez-Miller, S. J.; Hurley, T. D. Coenzyme isomerization is integral to catalysis in aldehyde dehydrogenase. *Biochemistry* **2003**, 42, 7100-9.

151. Lipinski, C.; Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **2004**, 432, 855-61.

152. Zheng, C. F.; Wang, T. T.; Weiner, H. Cloning and expression of the full-length cDNAs encoding human liver class 1 and class 2 aldehyde dehydrogenase. *Alcohol Clin Exp Res* **1993**, 17, 828-31.

153. Studier, F. W.; Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **1986**, 189, 113-30.

154. Krupenko, S. A. FDH: an aldehyde dehydrogenase fusion enzyme in folate metabolism. *Chem Biol Interact* **2009**, 178, 84-93.

155. Hammen, P. K.; Allali-Hassani, A.; Hallenga, K.; Hurley, T. D.; Weiner, H. Multiple conformations of NAD and NADH when bound to human cytosolic and mitochondrial aldehyde dehydrogenase. *Biochemistry* **2002**, 41, 7156-68.

156. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* **1997**, 276, 307-326.

157. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **1994**, 50, 760-3.

158. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004**, 60, 2126-32.

159. Painter, J.; Merritt, E. A. A molecular viewer for the analysis of TLS rigid-body motion in macromolecules. *Acta Crystallogr D Biol Crystallogr* **2005**, 61, 465-71.
160. Painter, J.; Merritt, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **2006**, 62, 439-50.
161. Cleland, W. W. The statistical analysis of enzyme kinetic data. *Adv Enzymol Relat Areas Mol Biol* **1967**, 29, 1-32.
162. Parsons, Z. D.; Gates, K. S. Redox regulation of protein tyrosine phosphatases: methods for kinetic analysis of covalent enzyme inactivation. *Methods Enzymol* **2013**, 528, 129-54.
163. Aldridge, W. N.; Reiner, E. *Enzyme inhibitors as substrates. Interactions of esterases with esters of organophosphorus and carbamic acids*. North-Holland Pub. Co.: Amsterdam, 1972; p xvi, 328 p.
164. Pappa, A.; Estey, T.; Manzer, R.; Brown, D.; Vasiliou, V. Human aldehyde dehydrogenase 3A1 (ALDH3A1): biochemical characterization and immunohistochemical localization in the cornea. *Biochem J* **2003**, 376, 615-23.
165. Marselos, M.; Lindahl, R. Substrate preference of a cytosolic aldehyde dehydrogenase inducible in rat liver by treatment with 3-methylcholanthrene. *Toxicol Appl Pharmacol* **1988**, 95, 339-45.
166. Parajuli, B.; Georgiadis, T. M.; Fishel, M. L.; Hurley, T. D. Development of Selective Inhibitors for Human Aldehyde Dehydrogenase 3A1 (ALDH3A1) for the Enhancement of Cyclophosphamide Cytotoxicity. *Chembiochem* **2014**.
167. Morgan, C. A.; Hurley, T. D. Development of a high-throughput in vitro assay to identify selective inhibitors for human ALDH1A1. *Chem Biol Interact* **2014**, In press.
168. Kim, Y. G.; Lee, S.; Kwon, O. S.; Park, S. Y.; Lee, S. J.; Park, B. J.; Kim, K. J. Redox-switch modulation of human SSADH by dynamic catalytic loop. *EMBO J* **2009**, 28, 959-68.
169. Pemberton, T. A.; Tanner, J. J. Structural basis of substrate selectivity of Delta(1)-pyrroline-5-carboxylate dehydrogenase (ALDH4A1): semialdehyde chain length. *Arch Biochem Biophys* **2013**, 538, 34-40.
170. MacGibbon, A. K.; Haylock, S. J.; Buckley, P. D.; Blackwell, L. F. Kinetic studies on the esterase activity of cytoplasmic sheep liver aldehyde dehydrogenase. *Biochem J* **1978**, 171, 533-8.

171. Lamb, A. L.; Newcomer, M. E. The structure of retinal dehydrogenase type II at 2.7 Å resolution: implications for retinal specificity. *Biochemistry* **1999**, *38*, 6003-11.
172. Ho, K. K.; Hurley, T. D.; Weiner, H. Selective alteration of the rate-limiting step in cytosolic aldehyde dehydrogenase through random mutagenesis. *Biochemistry* **2006**, *45*, 9445-53.
173. Lowe, E. D.; Gao, G. Y.; Johnson, L. N.; Keung, W. M. Structure of daidzin, a naturally occurring anti-alcohol-addiction agent, in complex with human mitochondrial aldehyde dehydrogenase. *J Med Chem* **2008**, *51*, 4482-7.
174. Berg, J. M.; Tymoczko, J. L.; Stryer, L.; Stryer, L. *Biochemistry*. 6th ed.; W.H. Freeman: New York, 2007.
175. Burleson, K. M.; Boente, M. P.; Pambuccian, S. E.; Skubitz, A. P. Disaggregation and invasion of ovarian carcinoma ascites spheroids. *J Transl Med* **2006**, *4*, 6.
176. Puiffe, M. L.; Le Page, C.; Filali-Mouhim, A.; Zietarska, M.; Ouellet, V.; Tonin, P. N.; Chevrette, M.; Provencher, D. M.; Mes-Masson, A. M. Characterization of ovarian cancer ascites on cell invasion, proliferation, spheroid formation, and gene expression in an in vitro model of epithelial ovarian cancer. *Neoplasia* **2007**, *9*, 820-9.

## Curriculum vitae

Cynthia A. Morgan

### Education

Indiana University, Indianapolis, IN  
Doctor of Philosophy in Biochemistry & Molecular Biology Jan 2015  
Thesis Advisor: Dr. Thomas D. Hurley  
Thesis Title: *Identification and characterization of small molecule inhibitors of aldehyde dehydrogenase 1A1 (ALDH1A1).*

Purdue University, Indianapolis, IN  
Master of Science in Mathematics 2001

Purdue University, Indianapolis, IN  
Bachelor of Science in Biology 1994  
Advisor: Dr. Martin Bard  
Thesis Title: *Cloning and characterization of the MPH gene, a morpholine sensitive gene in Saccharomyces cerevisiae*

### Research Experience

Indiana University School of Medicine 2010 – Current  
Department of Biochemistry & Molecular Biology  
Graduate Student Research Assistant: Enzyme kinetics and structural biology

- Performed *in vitro* high-throughput screen to identify novel, small molecule inhibitors of human ALDH1A1, an enzyme linked to cancer, obesity, and neurodegeneration
- Characterized compounds using enzyme kinetics, protein X-ray crystallography, protein expression and purification, site-directed mutagenesis, and cell culture

Indiana University School of Medicine 2000 – 2003, 2009 – 2010  
Institute of Psychiatric Research, Department of Psychiatry  
Research Technician: Six years research experience focused on the molecular biology of Alzheimer's disease and other neurodegenerative disorders

- Trained students and lab personnel; responsible for laboratory radiation safety program
- Laboratory skills include cell culture, enzyme assays, genotyping, protein and nucleic acid electrophoresis, and Western blotting; maintained inventory and ordered supplies

### Teaching Experience

Indiana University School of Medicine 2011

- Teaching Assistant – Biomedical Sciences I

Indiana University – Purdue University, Indianapolis 1993 – 1999

- Teaching Assistant: Zoology – Laboratory Section, 1995
- Teaching Assistant: Algebra – Discussion Section, 1994
- Academic Mentor: Led small discussion groups for first-year Calculus
- Tutored all levels of undergraduate mathematics

Indiana State Assistant Coach - Taekwondo 1995 – 2001

## Military Experience

United States Navy Reserve 2012 – Current

United States Air Force 2004 – 2009

### Effective Dates of Promotion

United States Air Force

Second Lieutenant 13 August 2004

First Lieutenant 13 August 2006

Captain 13 August 2008

United States Navy

Lieutenant 18 June 2012

## Awards, Grants, and Fellowships

Trainee Award, 17th International Workshop on the Enzymology and Molecular Biology of Carbonyl Metabolism, 2014

Best Presentation Award Co-Winner (IU Biochemistry Research Day), 2014

Kentucky Colonel by the Commonwealth of Kentucky, 2013

Jack Davis Award for Best Seminar by a Graduate Student, 2012 – 2013

Indiana University Graduate University Fellowship, 2010 – 2011

Air Force Commendation Medal, 2009

Meritorious Service Medal, 2008

Air Force Achievement Medal, 2006

Indiana University School of Medicine Fellowship in Aging, 2003

Indiana Sports Corporation Future Olympian Grant, 1997, 1998

United States Collegiate National Taekwondo Championship – Gold Medal, 1997

United States National Taekwondo Championship – Bronze Medal, 1996

## Publications

### Papers

Morgan CA and Hurley TD. *Characterization of two distinct structural classes of selective aldehyde dehydrogenase 1A1 inhibitors*. Submitted to Journal of Medicinal Chemistry.

Morgan CA, Parajuli B, Buchman C, Dria K, and Hurley TD. *N,N-diethylamino-benzaldehyde (DEAB) as a substrate and mechanism-based inhibitor for human ALDH isoenzymes*. Chemico-Biological Interactions. Epub: 12 Dec 2014.

Morgan CA and Hurley TD. *Development of a high-throughput in vitro assay to identify selective inhibitors for human ALDH1A1*. Chemico-Biological Interactions. Epub: 4 Nov 2014

Condello S., Morgan CA, Nagdas S, Cao L, Turek J, Hurley TD, and Matei D.  *$\beta$ -Catenin-regulated ALDH1A1 is a target in ovarian cancer spheroids*. Oncogene Epub: 23 June 2014.

Lahiri DK, Nall C, Chen D, Zaphiriou M, Morgan C, and Nurnberger Sr, JI. *Developmental expression of the  $\beta$ -amyloid precursor protein and heat shock protein-70 in the cerebral hemisphere region of the rat brain*. Ann. N.Y. Acad. Sci. 965, 324-333 (2002).



### **Abstracts / Posters / Presentations**

- Morgan CA and Hurley TD. *Identification and Characterization of Aldehyde Dehydrogenase 1A1 Selective Inhibitors*. Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, November 2014
- Morgan CA and Hurley TD. *Identification and Characterization of Aldehyde Dehydrogenase 1A1 Selective Inhibitors*. 17<sup>th</sup> International Meeting on Enzymology and Molecular Biology of Carbonyl Metabolism, Poconos, PA, July 2014.
- Morgan CA and Hurley TD. *Identification of small molecule inhibitors of aldehyde dehydrogenase 1A1 (ALDH1A1)*. Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, March 2012.
- Lahiri DK, Chen D, Morgan C, Utsuki T, Holloway HW, Ingram DK, and Greig NH. *Nicotine treatment increases the secretion of Alzheimer's  $\beta$ -amyloid precursor protein in cell culture and animal models*. At the ISN and ASN-sponsored satellite meeting entitled "Cellular and Molecular Mechanisms of Drugs of Abuse", Argentina, August 2001.
- Lahiri DK, Alley GM, Morgan C, Banerjee PK, and Farlow MR. *Effect of memantine on levels of the amyloid beta peptide in cell cultures*. 34th Annual meeting of the American Society for Neurochemistry, Newport Beach, CA, 2003. J. Neurochem.
- Lahiri DK, Morgan C, Chen D, Farlow MR, Sambamurti K, Rogers J, and Greig NH. *Cholinesterase inhibitors and amyloid-modifying properties*. 6th International Conference on Alzheimer's disease and Parkinson's disease, Seville, Spain, May 2003.
- Chen, D., Morgan, C., Yard, M., Kubek, M.J., Farlow, M., Rogers, J., Greig, N.H. and Lahiri, D.K. *Amyloid-Modifying Properties of Cholinesterase Inhibitors*. Soc. Neurosci. Abs. 29, 2003.