

RESEARCH ARTICLE

Diversity and Evolutionary Analysis of Iron-Containing (Type-III) Alcohol Dehydrogenases in Eukaryotes

Carlos Gaona-López, Adriana Julián-Sánchez, Héctor Riveros-Rosas*

Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México (UNAM). Cd. Universitaria, Ciudad de México, México

* hriveros@unam.mx



Abstract

Background

Alcohol dehydrogenase (ADH) activity is widely distributed in the three domains of life. Currently, there are three non-homologous NAD(P)⁺-dependent ADH families reported: Type I ADH comprises Zn-dependent ADHs; type II ADH comprises short-chain ADHs described first in *Drosophila*; and, type III ADH comprises iron-containing ADHs (FeADHs). These three families arose independently throughout evolution and possess different structures and mechanisms of reaction. While types I and II ADHs have been extensively studied, analyses about the evolution and diversity of (type III) FeADHs have not been published yet. Therefore in this work, a phylogenetic analysis of FeADHs was performed to get insights into the evolution of this protein family, as well as explore the diversity of FeADHs in eukaryotes.

Principal Findings

Results showed that FeADHs from eukaryotes are distributed in thirteen protein subfamilies, eight of them possessing protein sequences distributed in the three domains of life. Interestingly, none of these protein subfamilies possess protein sequences found simultaneously in animals, plants and fungi. Many FeADHs are activated by or contain Fe²⁺, but many others bind to a variety of metals, or even lack of metal cofactor. Animal FeADHs are found in just one protein subfamily, the hydroxyacid-oxoacid transhydrogenase (HOT) subfamily, which includes protein sequences widely distributed in fungi, but not in plants), and in several taxa from lower eukaryotes, bacteria and archaea. Fungi FeADHs are found mainly in two subfamilies: HOT and maleylacetate reductase (MAR), but some can be found also in other three different protein subfamilies. Plant FeADHs are found only in chlorophyta but not in higher plants, and are distributed in three different protein subfamilies.

Conclusions/Significance

FeADHs are a diverse and ancient protein family that shares a common 3D scaffold with a patchy distribution in eukaryotes. The majority of sequenced FeADHs from eukaryotes are

OPEN ACCESS

Citation: Gaona-López C, Julián-Sánchez A, Riveros-Rosas H (2016) Diversity and Evolutionary Analysis of Iron-Containing (Type-III) Alcohol Dehydrogenases in Eukaryotes. PLoS ONE 11(11): e0166851. doi:10.1371/journal.pone.0166851

Editor: Eugene A. Permyakov, Russian Academy of Medical Sciences, RUSSIAN FEDERATION

Received: August 25, 2016

Accepted: November 5, 2016

Published: November 28, 2016

Copyright: © 2016 Gaona-López et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by Dirección General de Asuntos del Personal Académico, Universidad Nacional Autónoma de México, PAPIIT grants IN216513 and IN225016 to HRR, and Consejo Nacional de Ciencia y Tecnología (México), Ph.D. fellowship No. 233791, and Programa de Doctorado en Ciencias, Universidad Nacional Autónoma de México, to CGL. The funders had no role in study design, data collection

and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

distributed in just two subfamilies, HOT and MAR (found mainly in animals and fungi). These two subfamilies comprise almost 85% of all sequenced FeADHs in eukaryotes.

1. Introduction

Alcohol dehydrogenase (ADH) activity is widely distributed in numerous phyla, which include organisms belonging to the three domains of life [1,2]. This activity is performed by different enzymes in different organisms. Indeed, there are three non-homologous NAD(P)⁺-dependent ADH families, which arose independently throughout evolution and possess different 3D scaffolds and mechanisms of reaction [3,4]. Type I ADHs were discovered first one hundred years ago by Federico Battelli and Lina Stern [5,6], who made the first preparation of a soluble alcohol dehydrogenase obtained from horse liver. Some years later, Bengt Andersson [7] showed that this enzyme requires the presence of co-enzyme or diphosphopyridine nucleotide (actually known as NAD⁺) to be active. In 1937, Erwin Negelein and Hans J. Wulff purified and crystallized an alcohol dehydrogenase from brewers' yeast [8], and in 1948, Roger K. Bonnichsen and Anders M. Wassen crystallized ADH from horse liver [9]. Few years later, Bert L. Vallee and Frederic L. Hoch showed that zinc is a functional component of the yeast and horse liver ADH [10,11]. Interestingly, horse liver ADH was also the first oligomeric enzyme for which an amino acid sequence [12] and a three-dimensional structure were determined [13].

In contrast, a type II ADH from *Drosophila melanogaster* was purified for the first time in 1968 by William Sofer and Heinrich Ursprung, who showed that this enzyme possesses a lower molecular weight as compared to that of liver and yeast ADHs, as well as a different substrate specificity [14]. Partial primary structure of *Drosophila* alcohol dehydrogenase obtained in 1976 [15] showed extensive differences with liver and yeast ADHs sequences, concluding that large differences exist between the active sites of the *Drosophila* enzyme and the other previously reported ADHs [15,16]. In 1981, Jörnvall and co-workers showed a distant but clear relationship among zinc-containing ADHs and sorbitol dehydrogenase from sheep, and between *Drosophila* ADH and ribitol dehydrogenase from *Klebsiella*, proposing that ADHs can be divided in "long chain" (type I) and "short chain" (type II) alcohol dehydrogenases [17].

A type III ADH was reported for first time by Christopher Wills and co-workers in 1981 [18] who found two ADHs with very different amino acid composition in *Zymomonas mobilis*. This new ADH-II was purified by Robert K. Scopes and described as an iron-activated ADH [19]. The gene which encodes this alcohol dehydrogenase II (*adhB*) from *Zymomonas mobilis* was cloned and sequenced by Tyrrell Conway and co-workers in 1987 [20] showing no homology with all previously sequenced ADHs. However, a few months later, Valerie M. Williamson and Charlotte E. Paquin [21] cloned a reported *ADH4* gene in *Saccharomyces cerevisiae* [22] showing that the amino acid sequence encoded by this *ADH4* gene was homolog to the iron-activated ADH II from *Z. mobilis*. A third homolog protein (1,2-propanediol oxidoreductase) encoded by *fucO* gene in *E. coli* was identified, allowing Tyrrell Conway and Lonnie O. Ingram to propose that these unusual ADHs comprise a novel (type III) ADH family of enzymes [23]. Later, new protein homologs to the iron-activated alcohol dehydrogenase (FeADH) family displaying different activities were found. Thus, glycerol dehydrogenase (GldA) from *Escherichia coli* [24]; butanol dehydrogenase (BdhA and BdhB) from *Clostridium acetobutylicum* [25]; ethanolamine utilization protein (EutG) from *Salmonella typhimurium* [26]; and, 1,3-propanediol dehydrogenase (DhaT) from *Klebsiella pneumoniae* [27] were all identified as homologs of the

FeADH family. Although type III ADHs were initially described only in microorganisms, Yingfeng Deng and co-workers identified and cloned, in 2002, a gene (*ADHFE1*) that encodes an iron-activated ADH in humans [28].

Nowadays, it has been shown that Zn-dependent (type I) ADHs are homologous to several other proteins that comprise the superfamily of medium-chain dehydrogenases/reductases (MDR) [2,29]; concurrently, short-chain (type II) ADHs belong to the superfamily of short-chain dehydrogenases/reductases (SDR), that comprise many different proteins with diverse catalytic and non-catalytic activities [30,31]. Oppositely, Iron-activated (type III) ADHs have not been extensively studied. According to NCBI's conserved domain database [32], iron-dependent ADHs are related to glycerol-1-phosphate dehydrogenases [33,34] and dehydroquinase synthases [34–36].

Several papers have been published analyzing the origin and evolution of Zn-dependent ADHs [37,38] and MDR superfamily [2], as well as the evolution of short-chain ADHs [1] and SDR superfamily [39]. However, analyses about the evolution and diversity of iron-activated (type III) ADHs have not been published yet. Therefore in this work, a phylogenetic analysis of iron-activated ADHs was performed, to get insights into the evolution of this protein family, as well as explore the diversity of iron-dependent ADHs in distinct eukaryotic phyla.

2. Methods

Amino acid sequences from eukaryotes belonging to FeADH family were retrieved by BlastP searches at the NCBI site [40] (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), or UniProt database [41] (<http://www.uniprot.org/>). Progressive multiple amino acid sequence alignments were performed with ClustalX version 2 [42] (<http://www.clustal.org/clustal2/>) using as a guide a structural alignment constructed with the VAST algorithm [43] (<http://www.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml>) that included all non-redundant Fe-ADHs protein structures deposited in the Protein Data Bank [44] (<http://www.rcsb.org/pdb/home/home.do>). Amino acid sequence alignments were corrected manually using BioEdit [45] (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>).

To obtain the smallest unbiased representative sample of protein sequences that are homologous to FeADHs, protein sequence dataset were collected from Pfam version 29.0 [46] based on representative proteomes [47] at 15% co-membership threshold (RP15). As FeADHs possess ca. 400 amino acids, only retrieved protein sequences with more than 200 residues were included in alignments and phylogenetic analyses.

Phylogenetic analyses were conducted using MEGA7 software [48] (<http://www.megasoftware.net>). Four methods were used to infer phylogenetic relationships: maximum likelihood (ML), maximum parsimony (MP), minimum evolution (ME), and neighborjoining (NJ). The amino acids substitution model described by Le-Gascuel [49], using a discrete Gamma distribution with five categories, was chosen as the best substitution model, since it gave the lowest Bayesian Information Criterion values and corrected Akaike Information Criterion values [50] in MEGA7 [48]. The gamma shape parameter value (+G parameter = 1.1824) was estimated directly from the data with MEGA7. Confidence for the internal branches of the phylogenetic tree, obtained using ML method, was determined through bootstrap analysis (500 replicates each).

Sequence logos were constructed using the WebLogo server (<http://weblogo.threeplusone.com/>). Each logo consists of stacks of amino acid letters. The ordinate axis of the logos graphs, indicate the stack for each position in the sequence. The height of the letters within the stack indicates the relative frequency of each amino acid at that position [51].

3. Results and Discussion

3.1. FeADH family definition

Iron-dependent (type-III) ADHs are reported as members of FeADH family in protein databases. However, public protein database use different criteria to sort amino acid sequences into different protein families and superfamilies; therefore, boundaries between related protein families are not necessarily the same. The NCBI's Conserved Domain Database [32] (<http://www.ncbi.nlm.nih.gov/cdd/>) identify iron-dependent (type-III) ADHs as members of DHQ-FeADH protein superfamily (cd07766), which comprises four related families: i) the dehydroquinase synthase-like family (cd08169), which catalyzes the conversion of 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) to dehydroquinase (DHQ) in the second step of the shikimate pathway; ii) the family of glycerol-1-phosphate dehydrogenase and related proteins (cd08549); iii) the glycerol dehydrogenase-like family (cd08550); and, iv) the iron-containing alcohol dehydrogenase-like family (cd08551). Pfam database [52] (<http://pfam.xfam.org/>) sorts these proteins into three different protein families: 1) the dehydroquinase synthase family (PF01761); 2) the iron-containing alcohol dehydrogenase family (PF00465); and, 3) the iron-containing alcohol dehydrogenase family 2 (PF13685).

To test the correspondence among the above described protein families, all identified sequences retrieved from the NCBI's Conserved Domain Database (152 sequences from cd08169 family; 66 sequences from cd08549 family; 118 sequences from cd08550; and 538 sequences from cd08551 family), were aligned with unbiased representative samples of protein sequences (15% co-membership threshold) collected from the Pfam families related with iron-containing ADHs (518 sequences from PF01761 family; 79 sequences from PF13685 family; and 1080 sequences from PF00465 family). Fig 1 shows an unrooted tree illustrating the correspondence between Pfam protein families and NCBI's Conserved Domain Database families. This figure shows that dehydroquinase synthase-like family (cd08169) shares the same branch as that protein sequences from PF01761 family in the Pfam database. In the same way, the family of glycerol-1-phosphate dehydrogenase and related proteins (cd08549) are located in the same branch as that the iron-containing alcohol dehydrogenase family 2 from Pfam database (PF13685). In contrast, the iron-containing alcohol dehydrogenase family (PF00465 from Pfam database) comprises amino acid sequences that belong to two related protein families in the NCBI's Conserved Domain Databases: the glycerol dehydrogenase-like family (cd08550); and the iron-containing alcohol dehydrogenase-like family (cd08551). Because glycerol dehydrogenases are reported as Zn-metallo-enzymes not containing iron [53,54], comprise a divergent branch with respect to the other iron-containing alcohol dehydrogenases (Fig 1), and conserve just one of the three conserved histidine residues involved in iron-binding (See 3.7 section), we centered the present analysis to the bona fide iron-dependent alcohol dehydrogenase (FeADH) protein family, as defined in the NCBI's Conserved Domain Database (cd08551).

3.2. (Type III) FeADH family comprises proteins with distinct catalytic activities

Several proteins reported as members from the FeADH family have been characterized exhibiting different catalytic activities. Thus, besides initial reports of iron-containing proteins with ethanol dehydrogenase activity in *Zymomonas mobilis* or *Saccharomyces cerevisiae* [18,21,23,55], other activities have been found: methanol dehydrogenase [56–58], lactaldehyde:propanediol oxidoreductase (lactaldehyde reductase) [23,59], propanol dehydrogenase [60], butanol dehydrogenase [61,62], L-1,3-propanediol dehydrogenase [63–66], maleylacetate

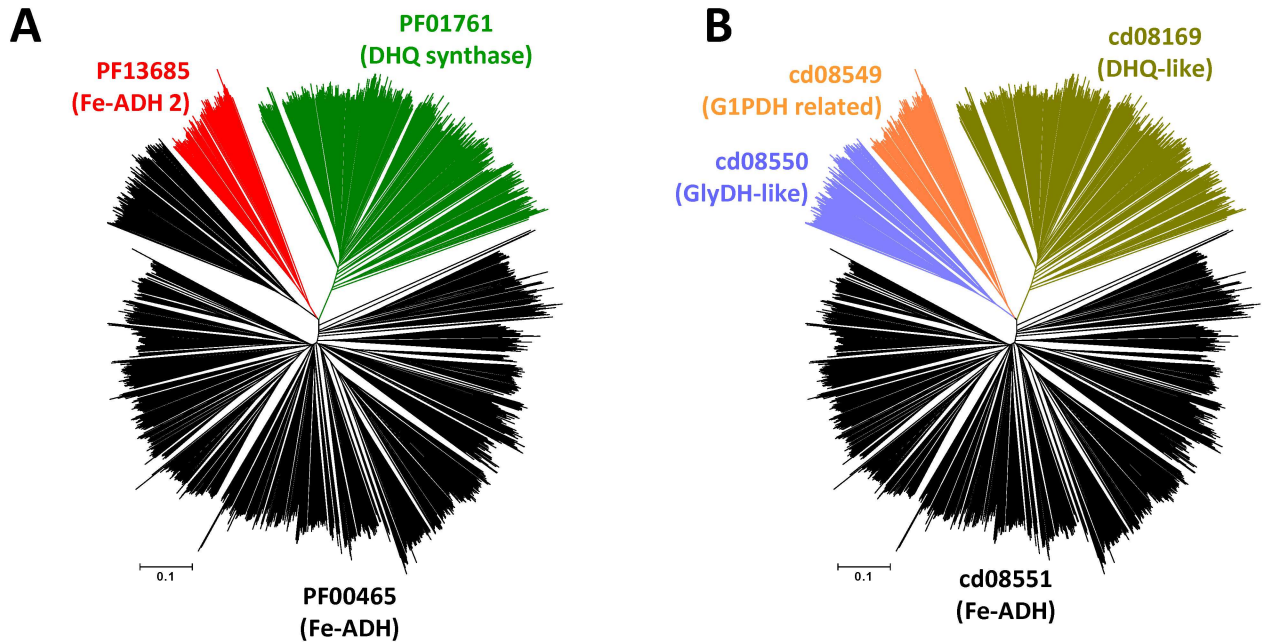


Fig 1. Unrooted tree constructed with protein sequences that possess homology to iron-dependent ADHs. 2459 nonredundant protein sequences were retrieved from Protein Data Bank, Swiss Prot database, NCBI's Conserved Domain Database, and Pfam database (using RP15 option to allow maximum representation of divergent proteins). Amino acid sequences were ascribed to protein families as considered by Pfam database (A) or NCBI's Conserved Domain Database (B).

doi:10.1371/journal.pone.0166851.g001

reductase [67–71], L-threonine dehydrogenase [72], and hydroxyacid-oxoacid transhydrogenase [73] among others.

3.3. (Type-III) FeADH family comprises several protein subfamilies

According to NCBI's Conserved Domain Database, sequences from FeADH protein family (cd08551) are distributed in at least 19 different protein subfamilies (Table 1). To explore the relationships between the different FeADH proteins, an alignment of 538 protein sequences retrieved from the NCBI's Conserved Domain Database, identified as members of any of the above mentioned 19 protein subfamilies was constructed, and used to perform a phylogenetic analysis. Fig 2 shows a maximum likelihood phylogenetic tree where it can be observed that each of the 19 protein subfamilies proposed by the NCBI's Conserved Domain Database possesses a good bootstrap support. Blast reciprocal best hits were used as an additional criterion (e.g., [74]) to corroborate that each of these families comprises a putative group of orthologous proteins (data not shown). On the other hand, among the different protein subfamilies comprised by the FeADH family, just a few closely related protein subfamilies showed a good bootstrap support between them. Thus, lactaldehyde:propanediol oxidoreductase (LPO) subfamily (cd08176) is related to FeADH4 subfamily (cd08188) (81% bootstrap support), and the C-terminal domain of the acetaldehyde-alcohol dehydrogenase two-domain (AAD-C) subfamily (cd08178) is related (95% bootstrap support) to butanol dehydrogenase (BDH) subfamily (cd08179) and propanediol dehydrogenase (PDD) subfamily (cd08180).

3.4. Phyletic distribution of FeADHs

Our results show that the FeADH family members are found in the three domains of life: archaea, bacteria, and eukarya (Table 1). In eukaryotes, FeADHs have a broad distribution and

Table 1. (Continued)

CDD Protein Subfamily	Reported activity [properties]/Characterized proteins [accession number]	Reported structure	Phyletic distribution			Reference
			Bacteria	Eukarya	Archaea	
AAAD-C cd08178	C-terminal alcohol dehydrogenase domain of the acetaldehyde dehydrogenase-alcohol dehydrogenase bifunctional two-domain protein [NAD(H)-dependent activity] ADHE Geobacillus thermoglucosidarius NCIMB 11955 [WP_013877698; C-terminal domain, 435–867 aa] [crystallized with Zn ²⁺] ADH2 <i>Ertamoeba histolytica</i> [Q24803] ADHE <i>Escherichia coli</i> /K-12 [P0A9Q7] Also possess activity as pyruvate-formatelyase deactivase	3ZDR — —	Yes	Yes (Amoebozoa; Alveolata; Diplomonadida; Cryptophyta; Viridiplantae ; Chlorophyta; Fungi , ascomycota, neocallimastigomycota)	No [80,100] [79–81] [82,83]	
NADPH-BDH cd08179	NADPH-dependent butanol dehydrogenase (<i>Clostridium saccharobutylicum</i> and <i>C. beijerinckii</i> use both ethanol and butanol as substrates) AdhA <i>Clostridium beijerinckii</i> /NRRL B592 [AAM18705] ADH1 <i>Clostridium saccharobutylicum</i> , formerly <i>C. acetobutylicum</i> [P13604]	— —	Yes	No	Yes [61] [62,149]	
PDD cd08180	1,3-propanediol dehydrogenase [NAD ⁺ -dependent activity] PduQ <i>Salmonella typhimurium</i> LT2 [Q9XDN0; NP_460997]	—	Yes	No	No [60]	
PDD-like cd08181	Putative 1,3-propanediol dehydrogenase-like [The enzyme bound NADP ⁺] TM0920 gene of <i>Thermotoga maritima</i> [Q9X022; WP_004080642] [crystallized with NADP ⁺ , Fe ³⁺ , or Zn ²⁺]	1O2D; 1VHD	Yes	Yes (Diplomonadida)	[85]	
HEPD cd08182	Hydroxyethylphosphoate dehydrogenase or phosphonoacetaldehyde reductase . Encoding gene is located inside an operon involved in the biosynthesis of phosphinothricin tripeptide (PTT), an antibiotic used as herbicide. pHpC <i>Streptomyces viridochromogenes</i> DSM 40736 [AAU00078] fomC <i>Streptomyces fradiae</i> [ACG70833]	— —	Yes	Yes ⁵ (Stramenopiles)	Yes [150,151] [152,153]	
BDH cd08187	Butanol dehydrogenase / aldehyde reductase NADP ⁺ -dependent activity with preference for alcohols longer than C3 YqjD <i>Escherichia coli</i> [Q46856] [crystallized with NADP ⁺ , Zn ²⁺] BDH (TM0820 locus) <i>Thermotoga maritima</i> MSB8 [NP_228629; Q9WZS7] [crystallized with NADP ⁺ , Fe ³⁺]	1OJ7; 4QGS 1VLJ	Yes	Yes (Stramenopiles; Amoebozoa; Parabasalidea)	[86] Unpublished	
HOT cd08190	Hydroxyacid-oxoacid transhydrogenase ADHFE1 <i>Homo sapiens</i> [Q81WW8] ADHFE1 <i>Rattus norvegicus</i> [Q4QQW3]	— —	Yes	Yes (Ichthyosporaea; Apusozoa; Stramenopiles; Amoebozoa; Rhizaria; Fungi ; Metazoa)	Yes [28,73,104,107,154]	
HVD cd08193	5-hydroxyvalerate dehydrogenase CpnD <i>Comamonas</i> sp. NCIMB 9872 [BAC22648]	—	Yes	Yes ⁴ (Haptophyceae; Viridiplantae , tracheophyta)	No [155]	
FeADH1 cd08185	None characterized	—	Yes	No	Yes	
FeADH2 cd08183	None characterized	—	Yes	Yes (Alveolata; Stramenopiles; Rhodophyta; Viridiplantae , Chlorophyta)	Yes ¹	
FeADH3 cd08184	3-deoxy-alpha-D-manno-octulosonate 8-oxidase . [Catalyzes the first step of the biosynthesis of Kdo8N (8-amino-3,8-dideoxy-D-manno-octulosonate), found in lipopolysaccharides of members of the <i>Shewanella</i> genus] KdnB <i>Shewanella oneidensis</i> [Q8EEB0]	—	Yes	No	No [156]	
FeADH4 cd08188	None characterized	—	Yes	No	Yes	

(Continued)

Table 1. (Continued)

CDD Protein Subfamily	Reported activity [properties]/[Characterized proteins [accession number]]	Reported structure	Phyletic distribution			Reference
			Bacteria	Eukarya	Archaea	
FeADH5 cd08189	None characterized		Yes	Yes (Euglenozoa)	No	
FeADH6 cd08194	None characterized		Yes	Yes (Alveolata; Haptophyceae; Rhizaria; Ichthyosporae; Fungi, chytridiomycota)	Yes	
FeADH8 cd08186	NADP⁺-dependent ADH , [oxidizes a series of primary aliphatic and aromatic alcohols, but shows a higher affinity for aldehyde substrates]. <i>Thermococcus paravinelliae</i> (formerly T. sp. ES1) [ACK56133] <i>Thermococcus</i> sp. AN1 [AAB63011] <i>Thermococcus hydrothermalis</i> [CAA74334]	— — —	Yes	Yes (Diplomonadida)	Yes	[87–90]
FeADH7 cd08192	Iron-containing alcohol dehydrogenases probably involved in the linear alkylbenzenesulfonate (LAS) degradation pathway [in <i>Parvibaculum lavamentivorans</i> the expression of the gene encoding this enzyme is induced during growth with LAS] ² . <i>Parvibaculum lavamentivorans</i> [ABS64400]	—	Yes	Yes	Yes ³	
HHD FeADH10 cd08191	6-hydroxyhexanoate dehydrogenase ChnD1 <i>Brevibacterium</i> sp. HCU [AAK73161]	—	Yes	No (Haptophyceae; Stramenopiles)	No	[157]

¹ Found in *Lokiarchaeum* sp. GC14_75.

² Personal communication from Dr. David Schleheck, University of Konstanz.

³ Found in *Thaumarchaeota* archaeon SCGC AB-539-E09.

⁴ Found in *Posidonia oceanica*, a Mediterranean seagrass.

⁵ Found in *Nannochloropsis gaditana*, an oleaginous microalgae.

doi:10.1371/journal.pone.0166851.t001

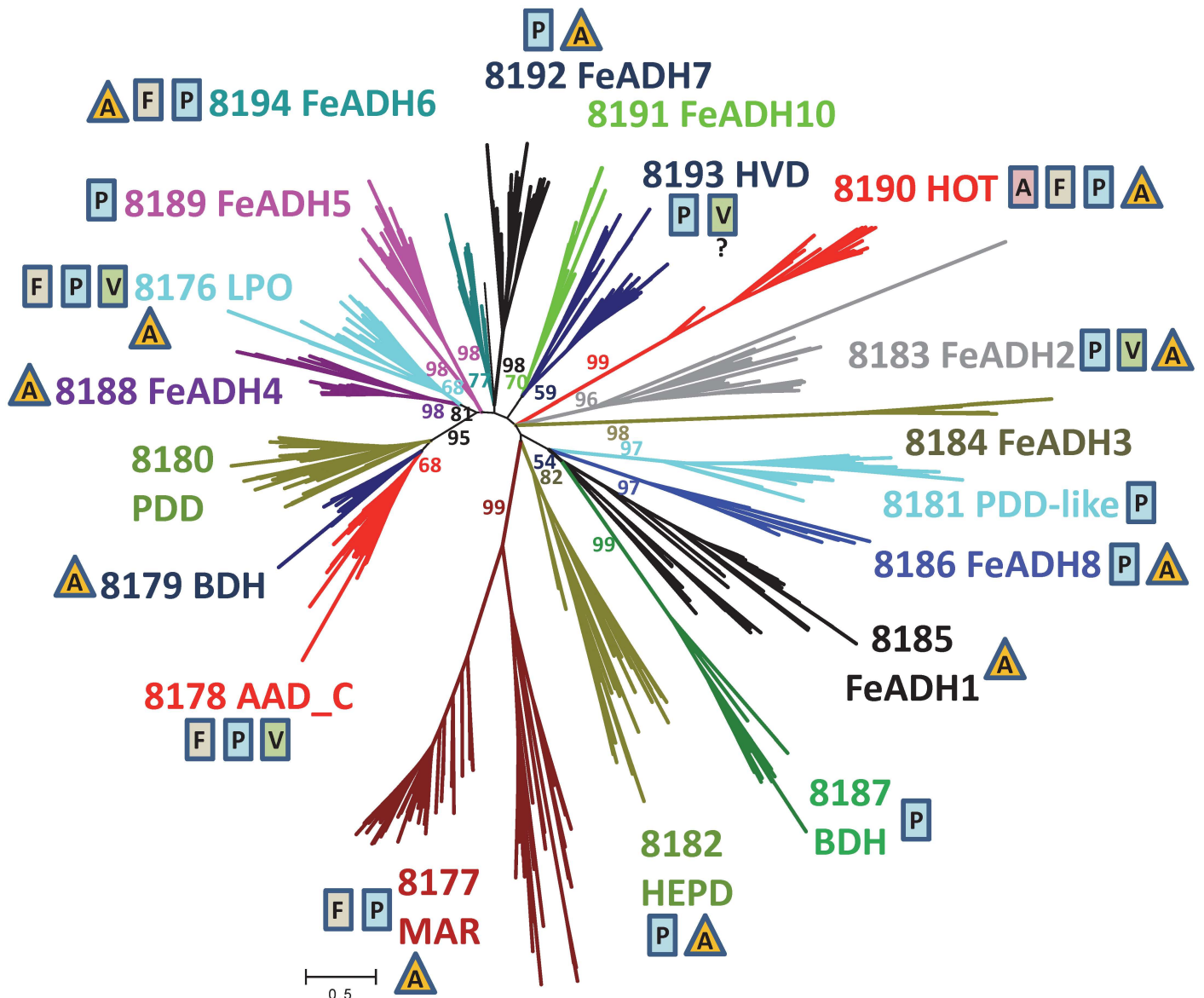


Fig 2. Phylogenetic analysis of 538 Fe-ADH protein sequences retrieved from the NCBI's Conserved Domain Database (CDD). The unrooted phylogenetic tree was inferred using the Maximum Likelihood method based on the Le-Gascuel model [49]. Branches are colored according to the Conserved Domain Database Fe-ADH subfamily they belong. The tree with the highest log likelihood (-2505413,5328) is shown. Similar trees were obtained with maximum-parsimony, minimum-evolution and neighbour-joining methods. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.8682)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. There were a total of 783 positions in the final dataset. The proportion of replicate trees in which the associated taxa clustered together in a bootstrap test (500 replicates) is given in color next to selected branches. Rectangles and triangles adjacent to each Fe-ADH subfamily name, indicate the presence of protein sequences from archaea domain (triangles), or eukarya domain (rectangles with A (animals), F (fungi), V (viridiplantae), and P (other eukaryotes) in each subfamily. Protein sequences from bacteria are present in all FeADH subfamilies.

doi:10.1371/journal.pone.0166851.g002

can be found in animals, fungi, plants and many lower eukaryotes. [S1 Table](#) provides a complete list of FeADH sequences from eukaryotes identified in this work. [Fig 3](#) shows a phylogenetic tree that comprises all identified FeADH subfamilies that possess proteins from eukaryotes. 656 protein sequences from eukaryotes (from a total of 868 sequences) are members of the HOT subfamily (cd08190). Thus, 75% of all sequenced eukaryotic FeADHs belongs to this subfamily. Indeed, all reported FeADH from animals (306 sequences), and 80% of

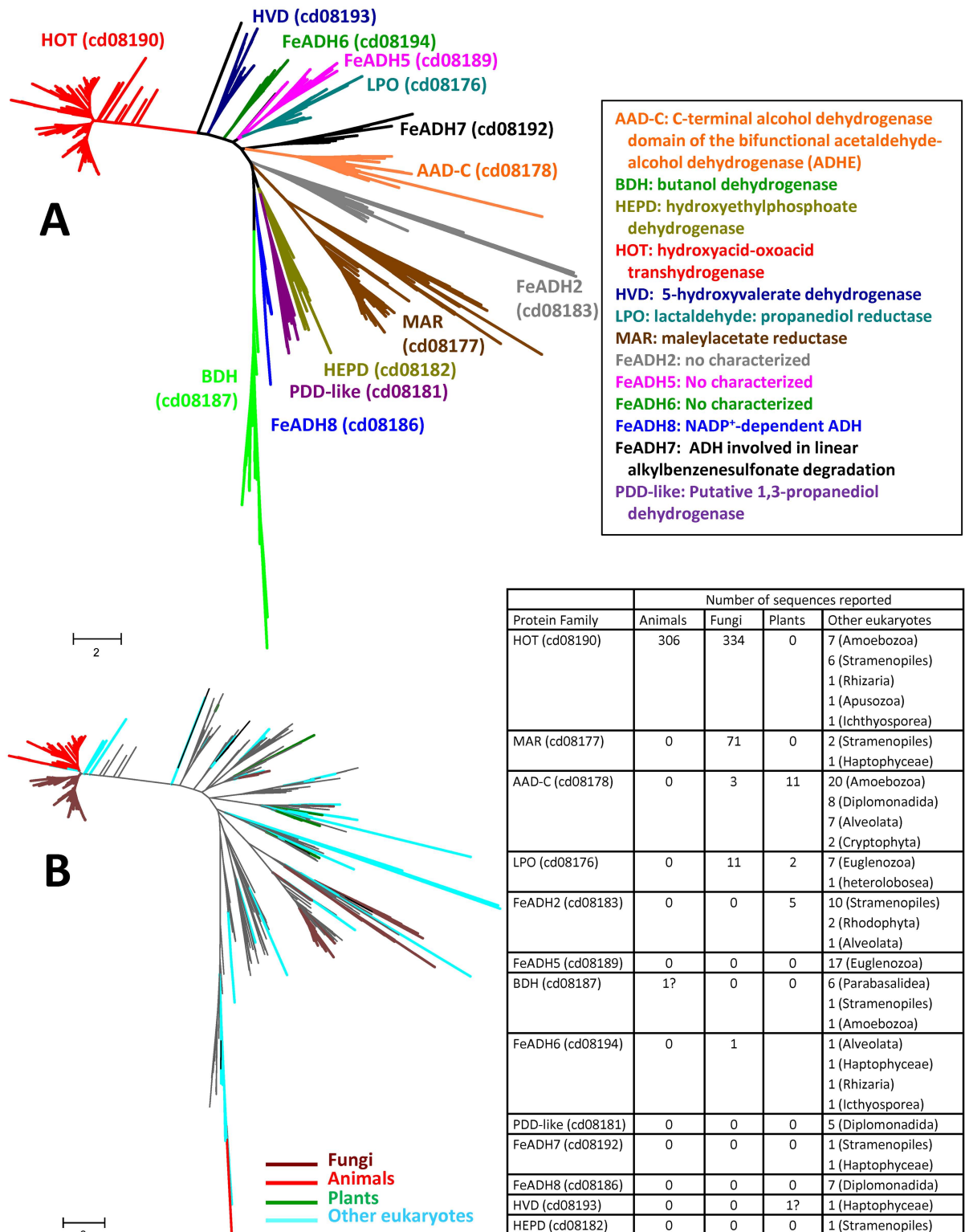


Fig 3. Phylogenetic analysis of 867 Fe-ADH protein sequences from eukaryotes plus 352 non-redundant sequences retrieved from the NCBI's Conserved Domain Database (CDD). The evolutionary history was inferred using the Maximum Likelihood method based on the Le-Gascuel model [1]. The tree with the highest log likelihood (-3414819.0869) is shown. Initial tree(s) for the heuristic search was/were obtained automatically applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete

Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.4901)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 1219 amino acid sequences. There were a total of 996 positions in the final dataset.

doi:10.1371/journal.pone.0166851.g003

FeADH found in fungi (334 sequences), belong to this protein subfamily. Other eukaryotes with HOT proteins are amoebozoa like *Acanthamoeba castellanii*, *Polysphondylium pallidum*, *Acytostelium subglobosum*, *Dictyostellium discoideum*, *D. lacteum*, and *D. purpureum*; stramenopiles like *Phaeodactylum tricornutum*, *Thalassiosira oceanica*, *Aphanomyces astaci*, *A. invadens*, *Saprolegnia parasitica* and *S. diclina*; ichthyosporea like *Capsaspora owczarzaki*; Apuzoa like *Thecamonas trahens*; and Rhizaria like the foraminifera *Reticulomyxa filose*. HOT sequences were not found in plants.

In fungi, FeADHs are sorted in two main protein subfamilies: 1) HOT subfamily (cd08190; that includes 80% of fungal protein sequences), is apparently found in all fungal taxa, with exception of some saccharomycetes which include yeast such as *Saccharomyces cerevisiae* and *Kluyveromyces lactis*, and schizosaccharomycetes such as *Schizosaccharomyces pombe*; and 2) MAR subfamily (cd08177), which includes almost 17% of fungal protein sequences, was found mainly in ascomycetes, and basidiomycetes. All reported FeADHs from saccharomycetes and schizosaccharomycetes belong to the LPO subfamily (cd08176) and probably are involved in ethanol metabolism (only the saccharomycete *Geotrichum candidum* was found to possess a FeADH that belong to the HOT subfamily). Three reported fungal sequences belong to the AAD-C (cd08178) subfamily (ADHE from *Togninia minima* (ascomycota), *Neocallimastix frontalis* and *Piromyces* sp. E2 (neocallimastigomycota)). However, the presence of ADHE in eukaryotes has been proposed to result from horizontal gene transfer from different bacteria [75]. In contrast, the FeADH from cd08194 subfamily found in *Gonapodya prolifera* JEL478 (chytridiomycota) is difficult to explain by horizontal gene transfer since the gene encoding this protein possesses 9 exons. Therefore, the origin of this last protein in fungi is uncertain.

FeADHs are absent in superior plants; only green algae (chlorophyta) possess FeADHs. Interestingly, different classes of green algae possess FeADH that belong to different protein subfamilies (Table 2). Thus, taxa from the chlorophyceae class possess only FeADH that belongs to the bidomain acetaldehyde-alcohol dehydrogenase (AAD-C) subfamily (cd08178). Algae's from class Trebouxiophyceae possess FeADHs that belongs to AAD-C (cd08178) and lactaldehyde: propanediol oxidoreductase (LPO) subfamily (cd08176); and algae's from the Class Prasinophyceae possess FeADHs that belong to LPO (cd08176) and one uncharacterized protein subfamily (cd08183). The broad distribution of FeADHs in chlorophyta and its absence in higher plants, suggests that genes encoding FeADHs were lost in the last common ancestor of terrestrial plants. Here, it should be mentioned that a FeADH from 5-hydroxyvalerate dehydrogenase (HVD) subfamily (cd08193) has been reported in the Mediterranean seagrass *Posidonia oceanica*. This protein was sequenced from an isolated mRNA that changed its expression in response to cadmium treatment [76], and showed the highest identity (65%) with FeADHs from *Rhizobium* genera (α -proteobacteria). It is not clear if this reported FeADH in *Posidonia oceanica* might results from horizontal gene transfer from a bacterium, or if it is just the results of bacterial contamination during the total RNA isolation procedure from leaves and apical tips. Because the presence of FeADHs has not been confirmed in any other higher plant, and the absence of additional evidence, its presence in *Posidonia oceanica* should be considered dubious.

It is interesting that fungi and chlorophyta exhibit a patchy distribution of FeADHs, particularly if it is considered that the FeADHs that belong to different protein subfamilies are not functionally equivalent and participate in different metabolic functions.

Table 2. Number of FeADH proteins from plants found in different subfamilies.

Organism	Number of genes			
	cd08178	cd08176	cd08183	cd08193
	AAD-C	LPO	FeADH2	HVD
Chlorophyta				
Class Chlorophyceae				
<i>Chlamydomonas reinhardtii</i> (CC-503 cw92 mt+)	3	0	0	0
<i>Polytomella</i> sp. Pringsheim 198.80	1	0	0	0
<i>Volvox carteri</i> f. Nagariensis	2	0	0	0
<i>Gonium pectoral</i>	2	0	0	0
<i>Monoraphidium neglectum</i> SAG 48.87	1			
Class Trebouxiophyceae				
<i>Chlorella variabilis</i> NC64A	2	0	0	0
<i>Auxenochlorella protothecoides</i> 0710 (<i>Chlorella protothecoides</i>)	0	1	0	0
Class Prasinophyceae				
<i>Micromonas pusilla</i> CCMP1545	0	1	1	0
<i>Micromonas</i> sp. RCC299 (<i>Micromonas commoda</i>)	0	0	1	0
<i>Ostreococcus tauri</i>	0	0	1	0
<i>Ostreococcus lucimarinus</i> CCE9901	0	0	1	0
<i>Bathycoccus prasinos</i>	0	0	1	0
Streptophyta, Tracheophyta (monocot)				
<i>Posidonia oceánica</i> (Mediterranean seagrass)	0	0	0	1?

doi:10.1371/journal.pone.0166851.t002

3.5. FeADHs share the same scaffold

The three-dimensional structures of twelve FeADH proteins have been resolved. These structures are sorted in five different protein subfamilies that belong to the FeADH family (cd08551). All FeADHs have two distinct domains separated by a deep cleft. The α/β N-terminal domain shows a Rossmann-fold structure and contains the coenzyme-binding site. The C-terminal domain is composed of nine α -helices and contains the iron-binding site. In Fig 4 it can be observed that this scaffold is conserved in all members of FeADH protein family as well as in members of related protein families such as glycerol dehydrogenase (GDH) family (cd08550), glycerol-1-phosphate (G1PDH) family (cd08549), and dehydroquinase synthase (DHQ) family (cd08169), which belong to the DHQ-FeADH protein superfamily (cd07766). However, in the DHQ family, the C-terminal domain comprises two or four β -strands in addition to the nine α -helices. The sequence identity among proteins that belong to different FeADH subfamilies is ca. 20% (30–40% sequence similarity), while the sequence identity among proteins that belong to different protein families inside the DHQ-FeADH protein superfamily (cd07766) is ca. 10% (20% similarity). Thus, although protein sequences from different subfamilies are divergent within the FeADH family, Fig 4 shows that they all share a similar scaffold and similar domains.

Figs 5 and 6 show a structure-based multiple sequence alignment of FeADHs with known 3D structure. It can be observed that the twenty-one secondary structures that exhibit the FeADH scaffold are strictly conserved in all FeADHs reported structures (eight β -strands and thirteen α -helices). The N-terminal domain comprises residues 1–229 (human ADHFE1 numbering), while the C-terminal domain comprises residues 230–467 (human ADHFE1 numbering) with the last nine α -helices. Thus, N-terminal domain is involved in binding the coenzyme NAD(P)H; and C-terminal domain possesses the conserved amino acids important for metal ion coordination. For comparative purposes, four structures of glycerol phosphate

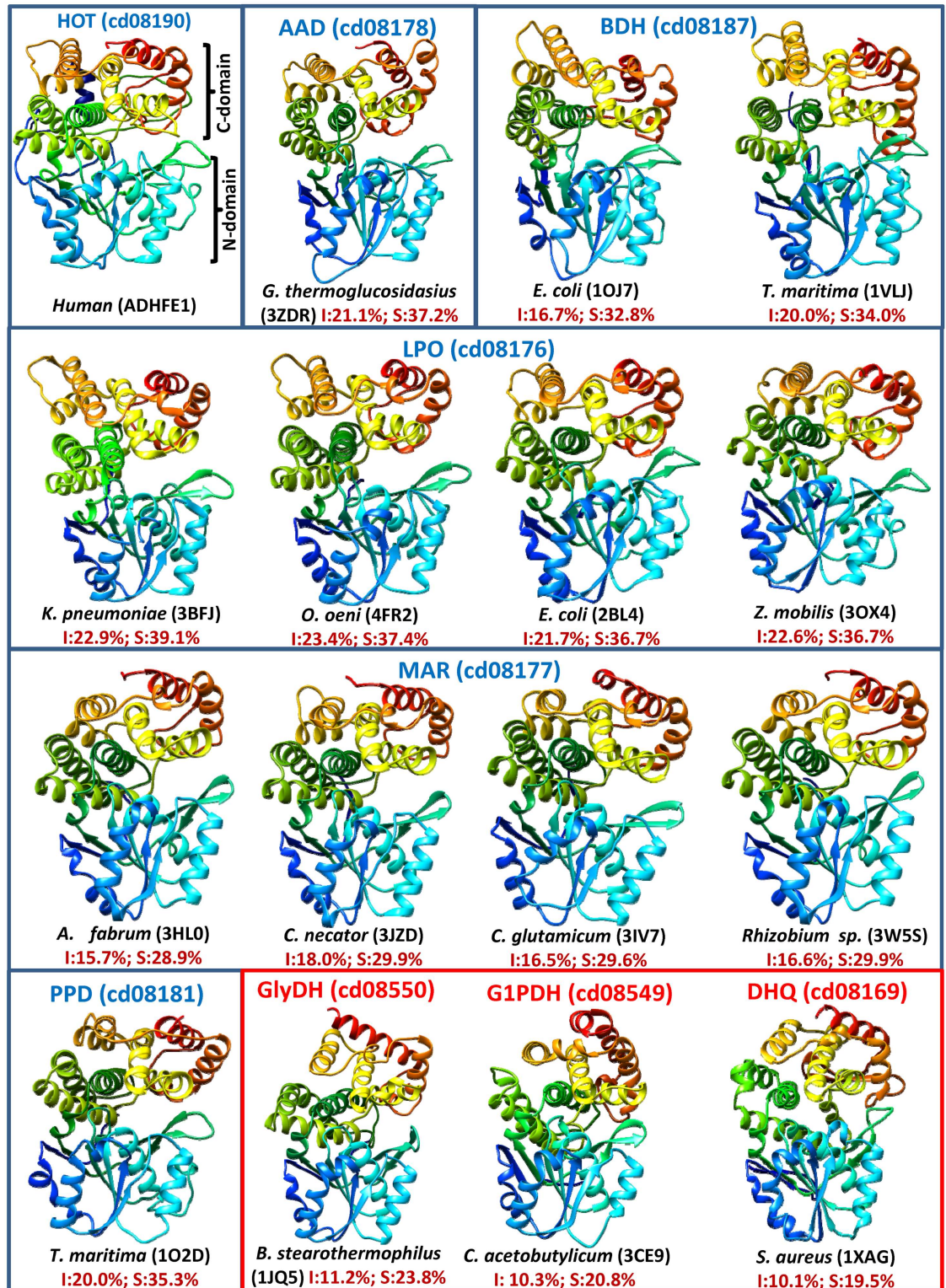


Fig 4. Comparison of the different FeADH proteins with a known three-dimensional structure. These proteins belong to five different FeADH protein subfamilies (sorted inside blue rectangles according to the protein subfamily to which they belong). Below each structure the scientific name of the organisms where the protein is found, as well as the PDB accession number is indicated in parenthesis. In a red rectangle are included representative structures of proteins of homolog protein families that belong to the DHQ-FeADH protein superfamily (cd07766). For reference, a structure

prediction (performed with I-TASSER server; [77]) of human ADHFE1 (accession NP_653251), which belongs to HOT subfamily (cd08190), is also included. Numbers in dark red show sequence identity (I) and similarity (S) between human ADHFE1 sequence and the indicated proteins. Secondary structure elements are colored in rainbow successive colors, starting from blue for the N-terminus and ending with red at the C-terminus. Protein structures were drawn using UCSF Chimera version 1.9 [78]

doi:10.1371/journal.pone.0166851.g004

from the related protein family cd08550 are included in Figs 5 and 6. They show a similar scaffold to FeADHs, but with some differences: the loop located between the $\alpha 4$ helix and the $\beta 5$ strand is very short in glycerol dehydrogenases; the $\alpha 6$ helix is displaced eight residues, and helices $\alpha 7$ and $\alpha 8$ are joined in one helix. All these differences, together with data from Fig 1, support the idea that glycerol dehydrogenases must be considered as a related protein family separated from *bona fide* FeADH family.

3.6. Coenzyme-binding site

As mentioned in the previous section, the N-terminal domain shows a Rossmann-type fold that contains the coenzyme-binding site. Residues involved in coenzyme binding are conserved in all FeADH subfamilies. A GGS motif (residues 138 to 141 according to human ADHFE1) is conserved in all FeADH subfamilies (Fig 5). This motif interacts with the pyrophosphate group of NAD(P)^+ and forms a loop that links the $\beta 4$ strand and the $\alpha 4$ helix.

Experimental support about coenzyme preference in FeADHs is scarce, but available data show that coenzyme specificity is mainly determined by the nature of the residue at position 81 (human ADHFE1 numbering): Eight different enzymes have been crystallized with their coenzyme as ligand: enzymes that bound NAD^+ possess aspartate or threonine at position 81, and enzymes that bound NADP^+ possess glycine at this position. Fig 7 shows the conservation of residues at position 81 as found in a logo analysis.

FeADHs with aspartate at position 81 prefer NAD^+ because the side-chain of this residue electrostatically and/or sterically repels the 2'-phosphate group of NADP^+ (the carboxyl group of Asp81 directly interacts with the hydroxyl group at C2 position of the adenine ribose). Examples are FucO from *E. coli* [23,59], DhaT from *Klebsiella pneumoniae* [66], ADH II from *Zymomonas mobilis* [18,23,55], MDH from *Bacillus methanolicus* [56–58], ADH4 from *Saccharomyces cerevisiae* [21,23], ADH2 from *Entamoeba histolytica* [79–81], and ADHE from *E. coli* [82,83].

In contrast, the shorter side-chain of glycine at position 81 is distant from the ribose and leaves room for binding the 2'-phosphate group of NADP^+ . Thus, enzymes having a residue with a shorter side-chain, can bind both, NAD^+ with less affinity, or NADP^+ , even sometimes with higher affinity than NAD^+ [84]. Examples of FeADHs with glycine at position 81 that bind NADP^+ are: 1,3-propanediol dehydrogenase (TM0920) from *Thermotoga maritima* [85], YqhD from *E. coli* [86], butanol dehydrogenase (TM0820) from *Thermotoga maritima* (PDB: 1VLJ), and FeADH from *Thermococcus hydrothermalis*, *T. paralvinellae* and *T. sp. AN1* [87–90].

Serine and threonine are two short-chain residues that have been associated in other NADP^+ -dependent enzymes, as residues that can bind to the 2'-phosphate group of NADP^+ [84,91]. However, HxqD from *Cupriavidus necator* JMP134, and MacA from *Agrobacterium fabrum*, are two enzymes that belong to the maleylacetate reductase subfamily (cd08177), which were crystallized with NAD^+ as ligand (PDB: 3JZD and 3HL0), and both possess threonine at position 81. In addition, maleylacetate reductase (Ncgl1112) from *Corynebacterium glutamicum* can use both coenzymes NAD^+ and NADP^+ [92], despite having a glycine at position 81. Therefore, although the residue at position 81 is the most important determinant of

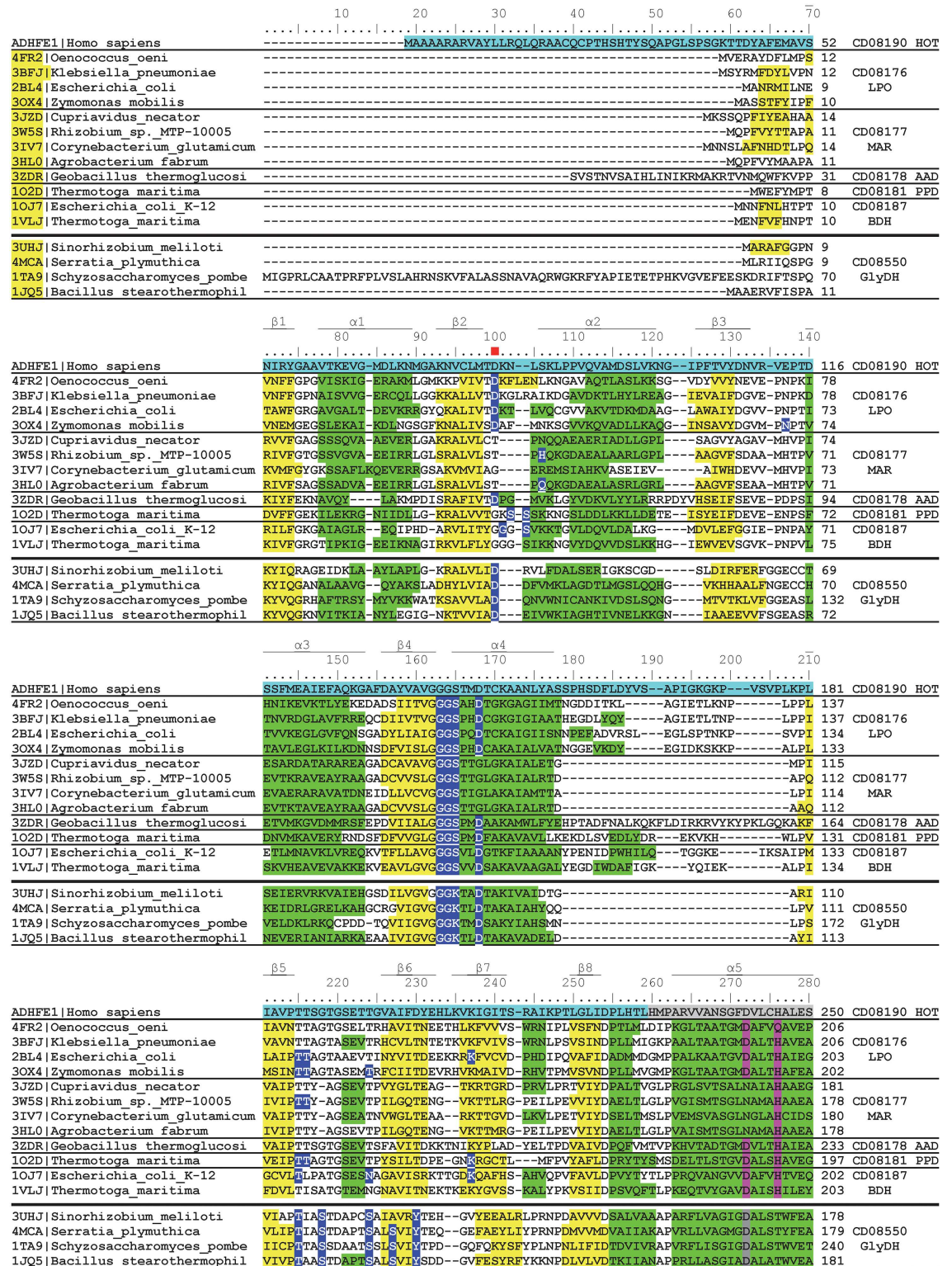


Fig 5. Multiple structure-based sequence alignment of FeADHs with a known 3D structure (residues 1–250 according to human ADHFE1). These proteins belong to five different subfamilies of the FeADH family. For comparison, ADHFE1 sequence from human is included in the alignment, as well as four glycerol dehydrogenase sequences with a known three-dimensional structure. PDB accession number of each sequence is indicated at the left side of alignment, whereas the protein subfamily to which each sequence belongs, is in the right side of the alignment. Conserved β-strands

and α -helices for each structure are indicated in yellow and green, respectively. Residue position determinant for coenzyme specificity is indicated with a red square. Residues involved in the binding of Fe atom are highlighted in pink; residues involved in the binding of Zinc atom in glicerol dehydrogenases are highlighted in grey. Amino acid residues from human ADHFE1 sequence, highlighted in blue and grey indicate positions that belong to the N-terminal or C-terminal domains, respectively. The three-dimensional alignment of FeADH structures was performed using the VAST tool at the NCBI's server [43].

doi:10.1371/journal.pone.0166851.g005

coenzyme preference, additional residues must be considered. A similar conclusion has been obtained in other NAD(P)-dependent enzymes as for example aldehyde dehydrogenases [84,93,94].

On the other hand, González-Segura et al. [84] analyzed the coenzyme preference of different aldehyde dehydrogenase (ALDH) families and found that coenzyme preference is a variable feature within many ALDH families, consistent with being mainly dependent on a single residue that apparently has no other structural or functional role, and therefore can easily be changed through evolution and selected in response to physiological needs. Considering that residues at position 81 are not conserved in some FeADH subfamilies (e.g., MAR subfamily (cd08177), hydroxyethylphosphoate dehydrogenase (HEPD) subfamily (cd08182), FeADH2 subfamily (cd08183) and FeADH8 subfamily (cd8186)), it is likely that in these subfamilies, coenzyme preference is a variable feature also.

3.7. Metal-binding site

The majority of FeADH subfamilies, contain a divalent metal M^{2+} , which is tetrahedrally coordinated through an ion dipole interaction with four conserved residues: Asp242, His246, His330, and His357 (according to human ADHFE1 numbering) (Figs 5–7). Interestingly, maleylacetate reductases (MAR subfamily; cd08177) are active in absence of metal ions, and do not have a divalent metal M^{2+} at their active center [69,95–97]; this may be due to the substitution of Asp242 by asparagine or arginine (Figs 5–7), which probably makes that MAR enzymes lose affinity for metal ions [95]. Considering this, members of the uncharacterized subfamily FeADH2 (cd08183), and some members of the HEPD subfamily (cd08182), probably are also functional in absence of divalent metal because Asp242 is replaced by glutamine in these proteins (Fig 7).

In a previous study using site-directed mutagenesis [98], the Fe^{2+} -binding participation of His267 from *E. coli* FucO was proposed (Tyr334 according to human ADHFE1). However, the crystal structure of *E. coli* FucO showed that His267 is not coordinated with Fe^{2+} ions [59]. Recently, Fujii et al., [95] performed a structural comparison between *E. coli* FucO and *Rhizobium* sp. MTP-10005 GraC (an enzyme with maleylacetate reductase activity; MAR subfamily; cd08177), and proposed that His267 of FucO correspond to His 243 of GraC, and that both residues could interact with the substrate, and therefore should be involved in catalysis, but not in metal-binding.

It is important to note that despite the members of the FeADH family (cd08551) are described as “iron-activated” alcohol dehydrogenases, these enzymes are activated by a range of divalent cations, among which, besides iron, we can find others such as zinc, nickel, magnesium, copper, cobalt, or manganese (e.g., [64,86,99,100]). Moreover, in enzymes activated by iron such as *E. coli* FucO or *Z. mobilis* ADH II (from LPO subfamily; cd08176), iron can be displaced by zinc [59,101]. *E. coli* FucO is an interesting example, because although FucO is active only with Fe^{2+} (Zn^{2+} inactivates the enzyme), FucO has *in vitro* a higher affinity for Zn^{2+} than for Fe^{2+} [59].

In the glycerol dehydrogenases, iron is absent but they contain a zinc-atom coordinated by two histidines and one aspartate [102,103]. Thus, only one histidine residue at position 357

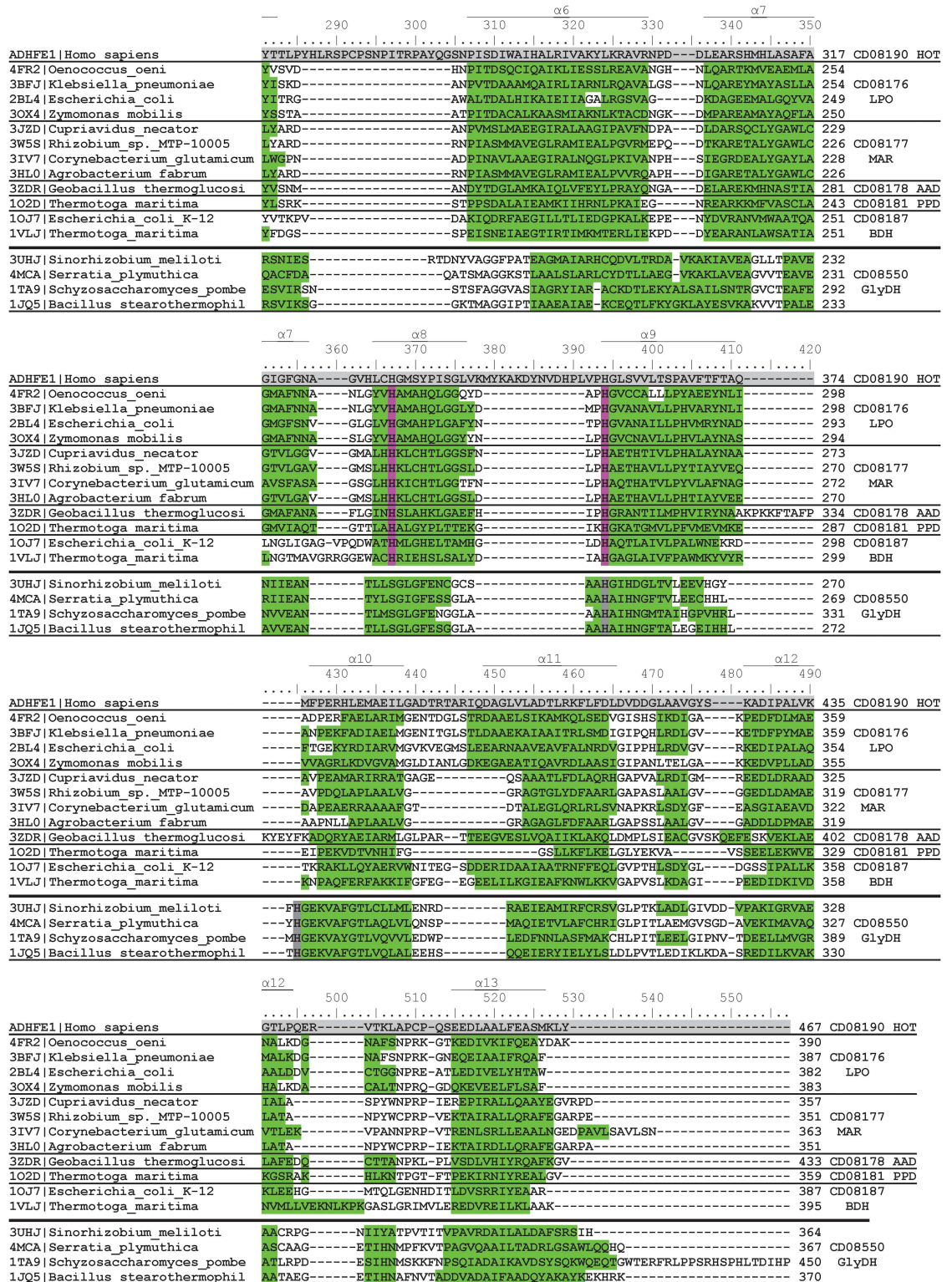


Fig 6. Multiple structure-based sequence alignment of FeADHs with a known 3D structure (residues 251–467 according to human ADHFE1). For additional details see caption of Fig 5.

doi:10.1371/journal.pone.0166851.g006

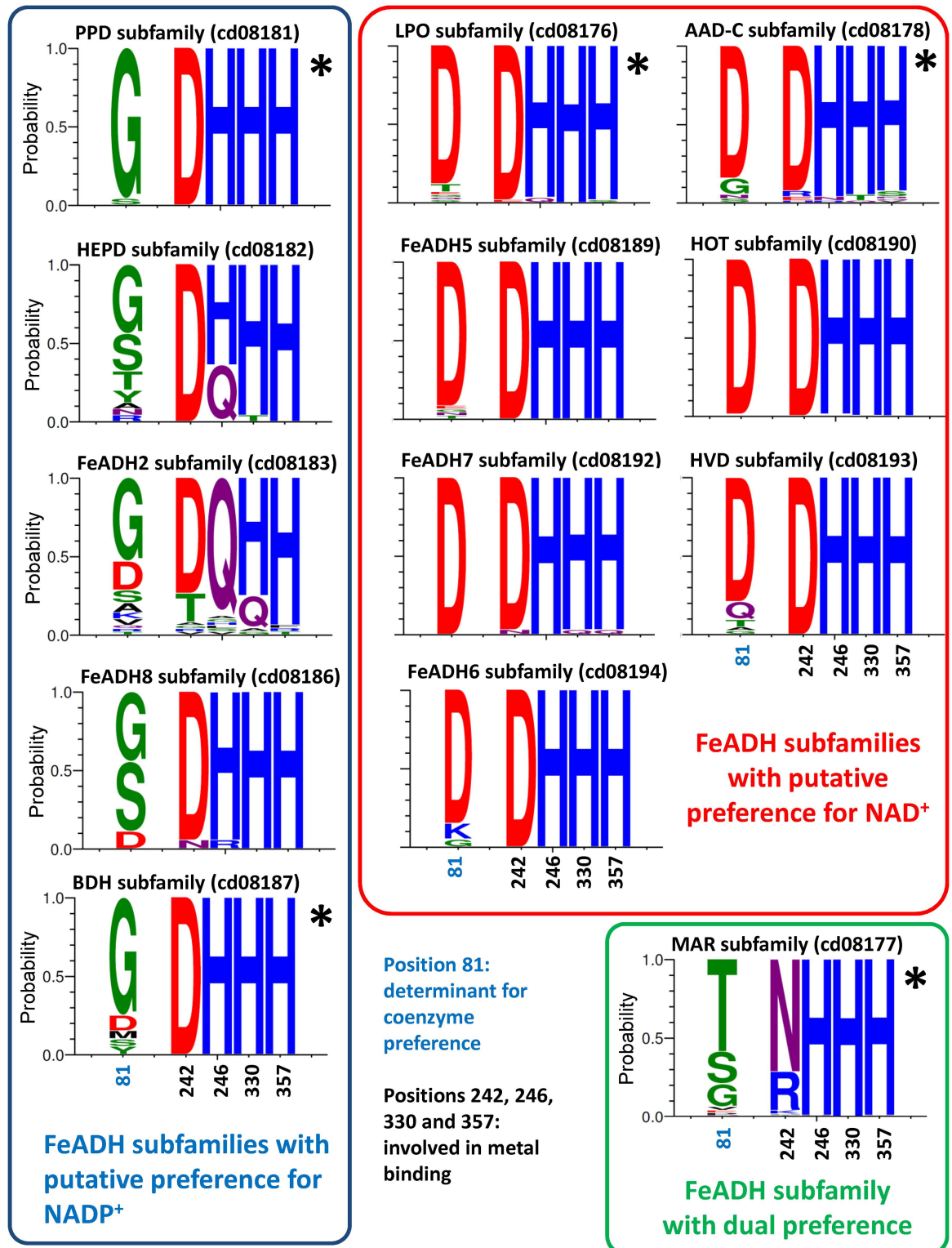


Fig 7. Sequence logos of selected positions in different FeADH subfamilies. The sequences of FeADH family were sorted in subfamilies according to the results of the phylogenetic analysis. Numbering is according to the sequence of human ADHFE1. Residue in position 81 is determinant for coenzyme preference; residues in positions 242, 246, 330 and 357 are involved in metal binding. The amino acid residue coloring scheme was according to their chemical properties: polar (G, S, T, Y, C), green; neutral (Q, N), purple; basic (K, R, H), blue; acidic (D, E), red; and hydrophobic (A, V, L, I, P, M, W, F), black.

FeADH subfamilies, whose members putatively use NADP⁺ as coenzyme, are enclosed with a blue box, those that use NAD⁺ as coenzyme, are enclosed with a red box, and those that use both NAD⁺ and NADP⁺, are enclosed in a green box. FeADH subfamilies with experimental support for coenzyme preference are indicated with an asterisk. Sequence logos were made using WebLogo 3 (<http://weblogo.threeplusone.com>) [51].

doi:10.1371/journal.pone.0166851.g007

(according to human ADHFE1 numbering) is conserved in both families, and is used to coordinate either an iron-atom in FeADHs, or a zinc-atom in glycerol dehydrogenases (see Figs 5 and 6). The differences in metal-binding residues between FeADHs and glycerol dehydrogenases support the idea that these latter proteins are members of a different, but related protein family.

3.8. Protein subfamilies that possess FeADHs from eukaryotes

All sequenced eukaryotic FeADHs are sorted in thirteen different protein subfamilies that belong to the FeADH family. Only one FeADH found in *Vitrella brassicaformis* CCMP3155 (Alveolata; Protein accession number: CEM34088) could not be ascribed to any of above identified protein subfamilies. However, because no Blast reciprocal best hits could be identified for this protein, we propose that this FeADH is just a divergent sequence and not a member of a new FeADH protein subfamily. Four of the FeADH subfamilies found in eukaryotes contain more than 92% of all FeADH sequences identified in these organisms. These subfamilies are:

3.8.1. HOT subfamily (cd08190). Some proteins of this subfamily have been characterized in mammals and possess activity as hydroxyacid-oxoacid transhydrogenase (HOT), catalyzing the conversion of γ -hydroxybutyrate into succinic semialdehyde in a reaction coupled with the reduction of α -cetoglutarate [73,104,105]. In humans, the gene encoding HOT was denominated *ADHFE1* [28] by the HUGO gene nomenclature committee. In animals, γ -hydroxybutyrate (GHB) is a naturally occurring compound present in micromolar concentration in brain and peripheral tissues [106], and HOT is the most active enzyme that oxidizes GHB [107]. GHB is of interest because it is a natural compound with neuromodulatory properties at central GABAergic synapses [108], is an energy regulator that promotes the release of growth hormone [109], and has been illegally used by athletes as a performance-enhancing drug [110]. Indeed, endogenous GHB metabolism appears to be associated with natural athletic ability [111]. This idea is supported by data that identify *ADHFE1* as an athletic-performance candidate gene, which has been a target for positive selection during 400 years in Thoroughbred horses [112].

The *ADHFE1* gene is expressed mainly in adult liver, kidney, heart, adipocytes [28,73,113], in hypothalamus and neuroblastoma cells [73], and diverse fetal tissues [28], as well as surface epithelium and crypt top of colorectal mucosa [114]. In contrast, *ADHFE1* transcript is non-detectable in lung, intestine, stomach, seminiferous tubules, muscle and testis [113]. Tae et al. [114] showed that *ADHFE1* expression in colon is higher in well-differentiated tissues than in poorly differentiated tissues, and that colorectal cancer cell lines show a down-regulation of *ADHFE1* mRNA and ADHFE1 protein due to hypermethylation of *ADHFE1* promoter. Therefore, ADHFE1 has an important role in organs with a high metabolic activity, as well as in differentiation and embryonic developmental processes.

Immunocytochemical staining reveals mitochondrial localization for mouse ADHFE1 [113]. Predictions performed with MITOPRED [115] WoLF PSORT [116], and PredSL [117], suggest that all animal ADHFE1s are also mitochondrial (data not shown). These enzymes conserve both the NAD(P)⁺-binding site and an iron-binding motif (see Figs 5–7). ADHFE1 contain a tightly bound cofactor and did not require the addition of NAD(P)⁺ to display catalytic activity [118]. Since ethanol oxidation requires coupling with the reduction of a second

molecule such as free NAD(P)⁺, participation of ADHFE1 on ethanol metabolism seems improbable. Furthermore, in human adipocytes exposure to ethanol (1–100 mM) does not modify the *ADHFE1* transcript levels [113], reinforcing the idea that this enzyme is not involved in ethanol metabolism in animals.

With no exceptions, all *ADHFE1* in animals corresponded to a single-copy gene, in spite of several whole genome duplications observed through the evolution of vertebrates. Because all animals have one *ADHFE1* that belongs to HOT subfamily, it can be assumed that this protein is performing essential activities in animals.

All HOT proteins possess two insertions: the first is a 19-residue insert (residues 256–274 in human *ADHFE1*) in a loop located between helices $\alpha 5$ and $\alpha 6$, and the second is a 13-residue insert (residues 342–354 in human *ADHFE1*) in a loop located between helices $\alpha 8$ and $\alpha 9$ (see Fig 6). This insert is absent in other iron-containing ADH members of FeADH family and even in the glycerol dehydrogenase protein family.

Because members of this protein subfamily are found in the three domains of life (archaea, bacteria and eukarya), this group is probably one of the most ancient protein subfamilies inside the FeADH family. However, the activity performed by this protein in non-animal organisms is unknown.

3.8.2. LPO subfamily (cd08176). This protein subfamily includes proteins with different catalytic activities (Table 1). Among the reported activities, we found: lactaldehyde:propanediol oxidoreductase (lactaldehyde reductase) [23,59], L-1,3-propanediol dehydrogenase [63–66], methanol dehydrogenase [56–58], alcohol dehydrogenase [18,21,23,55,119], and L-threonine dehydrogenase [72]. In eukarya, proteins that belong to this subfamily have been reported in fungi (saccharomycetes); chlorophyta (*Micromonas pusilla*), euglenozoa and heterolobosea. Of these proteins, only the *ADH4* from *Saccharomyces cerevisiae* has been thoroughly characterized [120–122]. *S. cerevisiae* possess five alcohol dehydrogenase (*Adh*) isoenzymes. Cultivation with glucose or ethanol as carbon substrate revealed that *ADH1* was the only alcohol dehydrogenase capable of efficiently catalyzing the reduction of acetaldehyde to ethanol [120]. A mutant yeast strain with the sole intact *ADH4* gene was able to grow on glucose but at much slower rates than the wild-type strains, to produce even less ethanol from glucose and was unable to utilize ethanol as carbon source [120]. In contrast, high levels of glycerol and acetaldehyde were observed in this mutant (op. cit.). Because *ADH4* transcription is not observed in strains grown on ethanol, and strains with *ADH4* as the only intact isoenzyme gene, were unable to grow on ethanol [120], it is likely that *ADH4* expression is not related to ethanol consumption, in spite of that the kinetic properties of *ADH4* compared with those of other yeast *ADHs* isoenzymes, showed that ethanol is a suitable substrate for *ADH4* [121,122]. Indeed, ethanol and n-propanol are the best substrates for yeast *ADH4* [121]. Thus, although the kinetic properties of *ADH4* make it suitable for ethanol metabolism, it is possible that this enzyme develops different physiological role(s).

3.8.3. AAD-C subfamily (cd08178). The C-terminal alcohol dehydrogenase domain of the bifunctional acetaldehyde dehydrogenase-alcohol dehydrogenase bidomain protein corresponds to one of the FeADH subfamilies found in bacteria, fungi, chlorophyta, and in several lower eukaryotes. These bifunctional bidomain enzymes are also known as ADHEs, and are found in many fermentative microorganisms. They catalyze the conversion of an acyl-coenzyme A to an alcohol via an aldehyde intermediate. This is coupled to the oxidation of two NADH molecules to maintain the NAD⁺ pool during fermentative metabolism. ADHE enzymes form large helical multimeric assemblies or ‘spirosomes’ [79,83,123], and consist of an N-terminal acetylating aldehyde dehydrogenase domain, which belongs to ALDH20 protein family of the ALDH superfamily, and a C-terminal alcohol dehydrogenase domain (*ADH*), which is a member of the AAD-C subfamily of the FeADH family.

ADHes have been described in many fermentative microorganisms that grow in anaerobic conditions, and it is generally accepted that ADHes perform the important function of regenerating NAD^+ from NADH under anaerobic conditions to maintain a continuous flow of glycolysis through alcoholic fermentation [80,124]. The fact that ADHE inhibition in *Entamoeba histolytica* induced a significant accumulation of glycolytic intermediates and lower ATP content [80], as well as the fact that ADHE knockout strains from bacteria and *E. histolytica* show the complete abolition of ethanol production and an inability to survive under anaerobic conditions [100,123], strongly support the role of ADHE in ethanol production. Indeed, the expression of the *adhE* gene is greatly increased under anaerobic conditions [81].

Atteia et al. [125] reported the presence of an ADHE in isolated mitochondria from the colorless chlorophyta *Polytomella* sp. Expression at ambient oxygen levels of ADHE in an oxygen-respiring algae extends the occurrence and expression of this enzyme to aerobic eukaryotes growing under aerobic conditions, and suggests that ADHE could be involved in either the maintenance of redox balance (ethanol production), or in ethanol assimilation (producing acetyl-CoA and NADH for respiration); and, depending upon environmental conditions, in both.

Finally, it is interesting to mention that *E. coli* ADHE can bind to 70S ribosome, exhibiting a substantial RNA unwinding activity, which can account for the ability of the ribosome to translate through downstream of at least certain mRNA helices [126]. Thus, ADHE can function in *E. coli* as a ribosomal regulatory protein, revealing an unexpected moonlighting action that opens the door to find additional functions in other ADHes.

3.8.4. Maleylacetate reductase (MAR) subfamily (cd08177). Proteins that belong to this subfamily have been described mainly as maleylacetate reductase (MAR), a key enzyme for degradation of ring-fission products derived from the aerobic microbial degradation of aromatic compounds [71]. They catalyze the NADH- or NADPH-dependent reduction of maleylacetate, at a carbon-carbon double bond, to 3-oxoadipate. We found MAR homologs in the three domains of life (Table 1). In eukaryotes, MARs are present mainly in fungi, including both ascomycetes and basidiomycetes. In lower eukaryotes, MAR homologs were found in Haptophyceae (*Emiliania huxleyi*) and stramenopiles (*Nannochloropsis gaditana*). In fungi, maleylacetate reductases contribute to the catabolism of very common substrates, such as tyrosine, resorcinol, phenol, hydroquinone, gentisate, benzoate, 4-hydroxybenzoate, protocatechuate, vanillate, and even, aromatic pollutants [96,127–132]. In *Fusarium verticilloides*, a MAR homolog gene identified as *FUM7* was found in a cluster of genes involved in fumonisin biosynthesis [133].

Fumonisin are polyketide mycotoxins that can accumulate in plants infected with this fungus and cause several fatal animal diseases, including leukoencephalomalacia in horses, pulmonary edema in swine, cancer in rats and mice, and esophageal cancer in humans [133,134].

FUM7-deletion mutants produce fumonisin analogs with an alkene function [135]. This suggests that *FUM7* likely catalyzes the reduction of an alkene intermediary of fumonisin biosynthesis, in a reaction similar to that performed by maleylacetate reductases.

4. Conclusions

FeADHs belong to an ancient protein family that can be found in the three domains of life. These proteins comprise a complex family with at least 19 different subfamilies with proteins that develop different metabolic functions. Many FeADHs are activated by or contain Fe^{2+} , but many others contain other divalent metals as Zn^{2+} , or even lack of metal cofactor. In eukarya, the majority of FeADHs belongs to the hydroxyacid oxoacid transhydrogenase (HOT) subfamily (cd08190). Indeed, 100% of FeADHs found in animals, and 80% of FeADHs

found in fungi, belong to this protein subfamily. Interestingly, HOT proteins are absent in plants. The rest of FeADHs from eukaryotes shows a patchy phyletic distribution, and are sorted in twelve additional protein families being the more important, the maleylreductase (MAR) subfamily (cd08177) found mainly in fungi, the lactaldehyde:propanediol dehydrogenase (LPO) subfamily (cd08176) and the bidomain aldehyde dehydrogenase-alcohol dehydrogenase (AAD) subfamily (cd08178) found in fungi, chlorophyta and lower eukaryotes. Several protein families with a patchy phyletic distribution have been reported previously, such as glucosamine-6-phosphate isomerase, alcohol dehydrogenase E, hybrid-cluster protein (*prisS*), A-type flavoprotein [75], glycerol-1-phosphate dehydrogenase [136], aerolysin [137], hemerythrin, hemocyanin, tyrosinase [138], phycocyanin-like phycobilisome proteins [139], and circularly permuted RAS-like GTPase domain [140] among others. The patchy distribution of these protein families has been explained mainly through intra- and inter-domain lateral gene transfer events, or gene transfer through endosymbiotic events in lower eukaryotes. Indeed, many genes of bacterial origin in eukaryotes were obtained through endosymbiotic events that generated actual mitochondria and chloroplast. Even more, many microbial eukaryotes obtained additional genes through secondary and tertiary eukaryote-eukaryote endosymbiosis events (e.g., [141–143]). This results in a very complex evolutionary history of lower eukaryotes and open the door to multiple events of gain/loss of protein genes and an extensive horizontal gene transfer. Thus, the scattered distribution of many FeADHs subfamilies in eukaryotes suggests that it is likely its presence/absence in different taxa results from events of lateral gene transfer or endosymbiotic gene transfer.

Supporting Information

S1 Table. Proteins identified in eukaryotes as members of different iron-containing alcohol dehydrogenase subfamilies.
(PDF)

Acknowledgments

We thank Dr. R. A. Muñoz-Clares (Fac. Química, UNAM) for critical reading of the manuscript and providing invaluable comments, and Drs. E. Piña (Fac. Medicina, UNAM), L.P. Martínez-Castilla (Fac. Química, UNAM) and L. Segovia-Forcella (Inst. Biotecnología, UNAM) for helpful discussions. This work is part of C. G-L. PhD thesis for Postgraduate Studies in Biomedical Sciences of UNAM (Doctorado en Ciencias Biomédicas), and was carried out at the School of Medicine with H. Riveros-Rosas as his PhD advisor. The authors want to thank Mrs. Josefina Bolado, Head of the Scientific Paper Translation Department, from División de Investigación at Facultad de Medicina, UNAM, for editing the English-language version of this manuscript.

Author Contributions

Conceptualization: CGL AJS HRR.

Data curation: CGL AJS HRR.

Formal analysis: CGL AJS HRR.

Funding acquisition: HRR CGL.

Investigation: CGL AJS HRR.

Methodology: CGL AJS HRR.

Project administration: HRR.

Resources: HRR.

Supervision: HRR.

Validation: CGL AJS HRR.

Visualization: CGL AJS HRR.

Writing – original draft: HRR.

Writing – review & editing: HRR CGL AJS.

References

1. Hernandez-Tobias A, Julian-Sanchez A, Piña E, Riveros-Rosas H (2011) Natural alcohol exposure: is ethanol the main substrate for alcohol dehydrogenases in animals? *Chem Biol Interact* 191: 14–25. doi: [10.1016/j.cbi.2011.02.008](https://doi.org/10.1016/j.cbi.2011.02.008) PMID: [21329681](https://pubmed.ncbi.nlm.nih.gov/21329681/)
2. Riveros-Rosas H, Julian-Sanchez A, Villalobos-Molina R, Pardo JP, Piña E (2003) Diversity, taxonomy and evolution of medium-chain dehydrogenase/reductase superfamily. *Eur J Biochem* 270: 3309–3334. PMID: [12899689](https://pubmed.ncbi.nlm.nih.gov/12899689/)
3. Jörnvall H, Landreh M, Ostberg LJ (2015) Alcohol dehydrogenase, SDR and MDR structural stages, present update and altered era. *Chem Biol Interact* 234: 75–79. doi: [10.1016/j.cbi.2014.10.017](https://doi.org/10.1016/j.cbi.2014.10.017) PMID: [25451589](https://pubmed.ncbi.nlm.nih.gov/25451589/)
4. Reid MF, Fewson CA (1994) Molecular characterization of microbial alcohol dehydrogenases. *Crit Rev Microbiol* 20: 13–56. doi: [10.3109/10408419409113545](https://doi.org/10.3109/10408419409113545) PMID: [8185833](https://pubmed.ncbi.nlm.nih.gov/8185833/)
5. Battelli F, Stern L (1909) L'alcoolase dans les tissus animaux. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris* 67: 419–421.
6. Battelli F, Stern L (1910) Die Alkoholoxydase in den Tiergeweben. *Biochem Z* 28: 145–168.
7. Andersson B (1934) Die co-Zymase als Co-Enzym bei enzymatischen dehydrierungen. *Hoppe-Seyler's Zeitschrift für physiologische Chemie* 225: 57–68.
8. Negelein E, Wulff HJ (1937) Diphosphopyridinproteid ackohol, acetaldehyd. *Biochem Z* 293: 351–389.
9. Bonnicksen RK, Wassen AM (1948) Crystalline alcohol dehydrogenase from horse liver. *Arch Biochem* 18: 361–363. PMID: [18875057](https://pubmed.ncbi.nlm.nih.gov/18875057/)
10. Vallee BL, Hoch FL (1955) Zinc, a component of yeast alcohol dehydrogenase. *Proc Natl Acad Sci U S A* 41: 327–338. PMID: [16589675](https://pubmed.ncbi.nlm.nih.gov/16589675/)
11. Vallee BL, Hoch FL (1957) Zinc in horse liver alcohol dehydrogenase. *J Biol Chem* 225: 185–195. PMID: [13416229](https://pubmed.ncbi.nlm.nih.gov/13416229/)
12. Jörnvall H (1970) Horse liver alcohol dehydrogenase. The primary structure of the protein chain of the ethanol-active isoenzyme. *Eur J Biochem* 16: 25–40. PMID: [4917230](https://pubmed.ncbi.nlm.nih.gov/4917230/)
13. Eklund H, Nordstrom B, Zeppezauer E, Soderlund G, Ohlsson I, Boiwe T, et al. (1976) Three-dimensional structure of horse liver alcohol dehydrogenase at 2–4 Å resolution. *J Mol Biol* 102: 27–59. PMID: [178875](https://pubmed.ncbi.nlm.nih.gov/178875/)
14. Sofer W, Ursprung H (1968) Drosophila alcohol dehydrogenase. Purification and partial characterization. *J Biol Chem* 243: 3110–3115. PMID: [5653193](https://pubmed.ncbi.nlm.nih.gov/5653193/)
15. Schwartz MF, Jörnvall H (1976) Structural analyses of mutant and wild-type alcohol dehydrogenases from drosophila melanogaster. *Eur J Biochem* 68: 159–168. PMID: [823018](https://pubmed.ncbi.nlm.nih.gov/823018/)
16. Jörnvall H (1977) Differences between alcohol dehydrogenases. Structural properties and evolutionary aspects. *Eur J Biochem* 72: 443–452. PMID: [320001](https://pubmed.ncbi.nlm.nih.gov/320001/)
17. Jörnvall H, Persson M, Jeffery J (1981) Alcohol and polyol dehydrogenases are both divided into two protein types, and structural properties cross-relate the different enzyme activities within each type. *Proc Natl Acad Sci U S A* 78: 4226–4230. PMID: [7027257](https://pubmed.ncbi.nlm.nih.gov/7027257/)
18. Wills C, Kratochil P, Londo D, Martin T (1981) Characterization of the two alcohol dehydrogenases of *Zymomonas mobilis*. *Arch Biochem Biophys* 210: 775–785. PMID: [7030207](https://pubmed.ncbi.nlm.nih.gov/7030207/)
19. Scopes RK (1983) An iron-activated alcohol dehydrogenase. *FEBS Lett* 156: 303–306. PMID: [6343121](https://pubmed.ncbi.nlm.nih.gov/6343121/)

20. Conway T, Sewell GW, Osman YA, Ingram LO (1987) Cloning and sequencing of the alcohol dehydrogenase II gene from *Zymomonas mobilis*. *J Bacteriol* 169: 2591–2597. PMID: [3584063](#)
21. Williamson VM, Paquin CE (1987) Homology of *Saccharomyces cerevisiae* ADH4 to an iron-activated alcohol dehydrogenase from *Zymomonas mobilis*. *Mol Gen Genet* 209: 374–381. PMID: [2823079](#)
22. Paquin CE, Williamson VM (1986) Ty insertions at two loci account for most of the spontaneous anti-mycin A resistance mutations during growth at 15 degrees C of *Saccharomyces cerevisiae* strains lacking ADH1. *Mol Cell Biol* 6: 70–79. PMID: [3023838](#)
23. Conway T, Ingram LO (1989) Similarity of *Escherichia coli* propanediol oxidoreductase (fucO product) and an unusual alcohol dehydrogenase from *Zymomonas mobilis* and *Saccharomyces cerevisiae*. *J Bacteriol* 171: 3754–3759. PMID: [2661535](#)
24. Truniger V, Boos W (1994) Mapping and cloning of *gldA*, the structural gene of the *Escherichia coli* glycerol dehydrogenase. *J Bacteriol* 176: 1796–1800. PMID: [8132480](#)
25. Walter KA, Bennett GN, Papoutsakis ET (1992) Molecular characterization of two *Clostridium acetobutylicum* ATCC 824 butanol dehydrogenase isozyme genes. *J Bacteriol* 174: 7149–7158. PMID: [1385386](#)
26. Stojiljkovic I, Baumler AJ, Heffron F (1995) Ethanolamine utilization in *Salmonella typhimurium*: nucleotide sequence, protein expression, and mutational analysis of the *cchA cchB eutE eutJ eutG eutH* gene cluster. *J Bacteriol* 177: 1357–1366. PMID: [7868611](#)
27. Yuanyuan Z, Yang C, Baishan F (2004) Cloning and sequence analysis of the *dhaT* gene of the 1,3-propanediol regulon from *Klebsiella pneumoniae*. *Biotechnol Lett* 26: 251–255. PMID: [15049372](#)
28. Deng Y, Wang Z, Gu S, Ji C, Ying K, Xie Y, et al. (2002) Cloning and characterization of a novel human alcohol dehydrogenase gene (*ADHFe1*). *DNA Seq* 13: 301–306. PMID: [12592711](#)
29. Persson B, Zigler JS Jr., Jornvall H (1994) A super-family of medium-chain dehydrogenases/reductases (MDR). Sub-lines including zeta-crystallin, alcohol and polyol dehydrogenases, quinone oxidoreductase enoyl reductases, VAT-1 and other proteins. *Eur J Biochem* 226: 15–22. PMID: [7957243](#)
30. Jörnvall H, Höög JO, Persson B (1999) SDR and MDR: completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS Lett* 445: 261–264. PMID: [10094468](#)
31. Persson B, Kallberg Y (2013) Classification and nomenclature of the superfamily of short-chain dehydrogenases/reductases (SDRs). *Chem Biol Interact* 202: 111–115. doi: [10.1016/j.cbi.2012.11.009](#) PMID: [23200746](#)
32. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43: D222–D226. doi: [10.1093/nar/gku1221](#) PMID: [25414356](#)
33. Guldán H, Sterner R, Babinger P (2008) Identification and characterization of a bacterial glycerol-1-phosphate dehydrogenase: Ni(2+)-dependent AraM from *Bacillus subtilis*. *Biochemistry* 47: 7376–7384. doi: [10.1021/bi8005779](#) PMID: [18558723](#)
34. Han JS, Kosugi Y, Ishida H, Ishikawa K (2002) Kinetic study of sn-glycerol-1-phosphate dehydrogenase from the aerobic hyperthermophilic archaeon, *Aeropyrum pernix* K1. *Eur J Biochem* 269: 969–976. PMID: [11846799](#)
35. Nichols CE, Ren J, Leslie K, Dhaliwal B, Lockyer M, Charles I, et al. (2004) Comparison of ligand-induced conformational changes and domain closure mechanisms, between prokaryotic and eukaryotic dehydroquininate synthases. *J Mol Biol* 343: 533–546. doi: [10.1016/j.jmb.2004.08.039](#) PMID: [15465043](#)
36. Sugahara M, Nodake Y, Sugahara M, Kunishima N (2005) Crystal structure of dehydroquininate synthase from *Thermus thermophilus* HB8 showing functional importance of the dimeric state. *Proteins* 58: 249–252. doi: [10.1002/prot.20281](#) PMID: [15508124](#)
37. Borrás E, Albalat R, Duester G, Pares X, Farres J (2014) The *Xenopus* alcohol dehydrogenase gene family: characterization and comparative analysis incorporating amphibian and reptilian genomes. *BMC Genomics* 15: 216. doi: [10.1186/1471-2164-15-216](#) PMID: [24649825](#)
38. Gonzalez-Duarte R, Albalat R (2005) Merging protein, gene and genomic data: the evolution of the MDR-ADH family. *Heredity (Edinb)* 95: 184–197.
39. Jörnvall H, Hedlund J, Bergman T, Oppermann U, Persson B (2010) Superfamilies SDR and MDR: from early ancestry to present forms. Emergence of three lines, a Zn-metalloenzyme, and distinct variabilities. *Biochem Biophys Res Commun* 396: 125–130. doi: [10.1016/j.bbrc.2010.03.094](#) PMID: [20494124](#)
40. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, et al. (2013) BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* 41: W29–W33. doi: [10.1093/nar/gkt282](#) PMID: [23609542](#)

41. Pundir S, Martin MJ, O'Donovan C (2016) UniProt Tools. *Curr Protoc Bioinformatics* 53: 1.
42. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948. doi: [10.1093/bioinformatics/btm404](https://doi.org/10.1093/bioinformatics/btm404) PMID: [17846036](https://pubmed.ncbi.nlm.nih.gov/17846036/)
43. Madej T, Lanczycki CJ, Zhang D, Thiessen PA, Geer RC, Marchler-Bauer A, et al. (2014) MMDB and VAST+: tracking structural similarities between macromolecular complexes. *Nucleic Acids Res* 42: D297–D303. doi: [10.1093/nar/gkt1208](https://doi.org/10.1093/nar/gkt1208) PMID: [24319143](https://pubmed.ncbi.nlm.nih.gov/24319143/)
44. Rose PW, Prlc A, Bi C, Bluhm WF, Christie CH, Dutta S, et al. (2015) The RCSB Protein Data Bank: views of structural biology for basic and applied research and education. *Nucleic Acids Res* 43: D345–D356. doi: [10.1093/nar/gku1214](https://doi.org/10.1093/nar/gku1214) PMID: [25428375](https://pubmed.ncbi.nlm.nih.gov/25428375/)
45. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41: 95–98.
46. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. (2014) Pfam: the protein families database. *Nucleic Acids Res* 42: D222–D230. doi: [10.1093/nar/gkt1223](https://doi.org/10.1093/nar/gkt1223) PMID: [24288371](https://pubmed.ncbi.nlm.nih.gov/24288371/)
47. Chen C, Natale DA, Finn RD, Huang H, Zhang J, Wu CH, et al. (2011) Representative proteomes: a stable, scalable and unbiased proteome set for sequence analysis and functional annotation. *PLoS One* 6: e18910. doi: [10.1371/journal.pone.0018910](https://doi.org/10.1371/journal.pone.0018910) PMID: [21556138](https://pubmed.ncbi.nlm.nih.gov/21556138/)
48. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*.
49. Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. *Mol Biol Evol* 25: 1307–1320. doi: [10.1093/molbev/msn067](https://doi.org/10.1093/molbev/msn067) PMID: [18367465](https://pubmed.ncbi.nlm.nih.gov/18367465/)
50. Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Syst Biol* 53: 793–808. doi: [10.1080/10635150490522304](https://doi.org/10.1080/10635150490522304) PMID: [15545256](https://pubmed.ncbi.nlm.nih.gov/15545256/)
51. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14: 1188–1190. doi: [10.1101/gr.849004](https://doi.org/10.1101/gr.849004) PMID: [15173120](https://pubmed.ncbi.nlm.nih.gov/15173120/)
52. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44: D279–D285. doi: [10.1093/nar/gkv1344](https://doi.org/10.1093/nar/gkv1344) PMID: [26673716](https://pubmed.ncbi.nlm.nih.gov/26673716/)
53. Spencer P, Bown KJ, Scawen MD, Atkinson T, Gore MG (1989) Isolation and characterisation of the glycerol dehydrogenase from *Bacillus stearothermophilus*. *Biochim Biophys Acta* 994: 270–279. PMID: [2493267](https://pubmed.ncbi.nlm.nih.gov/2493267/)
54. Spencer P, Slade A, Atkinson T, Gore MG (1990) Studies on the interactions of glycerol dehydrogenase from *Bacillus stearothermophilus* with Zn²⁺ ions and NADH. *Biochim Biophys Acta* 1040: 130–133. PMID: [2378897](https://pubmed.ncbi.nlm.nih.gov/2378897/)
55. Moon JH, Lee HJ, Park SY, Song JM, Park MY, Park HM, et al. (2011) Structures of iron-dependent alcohol dehydrogenase 2 from *Zymomonas mobilis* ZM4 with and without NAD⁺ cofactor. *J Mol Biol* 407: 413–424. doi: [10.1016/j.jmb.2011.01.045](https://doi.org/10.1016/j.jmb.2011.01.045) PMID: [21295587](https://pubmed.ncbi.nlm.nih.gov/21295587/)
56. Arfman N, Hektor HJ, Bystrykh LV, Govorukhina NI, Dijkhuizen L, Frank J (1997) Properties of an NAD(H)-containing methanol dehydrogenase and its activator protein from *Bacillus methanolicus*. *Eur J Biochem* 244: 426–433. PMID: [9119008](https://pubmed.ncbi.nlm.nih.gov/9119008/)
57. de Vries GE, Arfman N, Terpstra P, Dijkhuizen L (1992) Cloning, expression, and sequence analysis of the *Bacillus methanolicus* C1 methanol dehydrogenase gene. *J Bacteriol* 174: 5346–5353. PMID: [1644761](https://pubmed.ncbi.nlm.nih.gov/1644761/)
58. Vonck J, Arfman N, de Vries GE, Van BJ, van Bruggen EF, Dijkhuizen L (1991) Electron microscopic analysis and biochemical characterization of a novel methanol dehydrogenase from the thermotolerant *Bacillus* sp. C1. *J Biol Chem* 266: 3949–3954. PMID: [1995642](https://pubmed.ncbi.nlm.nih.gov/1995642/)
59. Montella C, Bellolell L, Perez-Luque R, Badia J, Baldoma L, Coll M, et al. (2005) Crystal structure of an iron-dependent group III dehydrogenase that interconverts L-lactaldehyde and L-1,2-propanediol in *Escherichia coli*. *J Bacteriol* 187: 4957–4966. doi: [10.1128/JB.187.14.4957-4966.2005](https://doi.org/10.1128/JB.187.14.4957-4966.2005) PMID: [15995211](https://pubmed.ncbi.nlm.nih.gov/15995211/)
60. Cheng S, Fan C, Sinha S, Bobik TA (2012) The PduQ enzyme is an alcohol dehydrogenase used to recycle NAD⁺ internally within the Pdu microcompartment of *Salmonella enterica*. *PLoS One* 7: e47144. doi: [10.1371/journal.pone.0047144](https://doi.org/10.1371/journal.pone.0047144) PMID: [23077559](https://pubmed.ncbi.nlm.nih.gov/23077559/)
61. Hiu SF, Zhu CX, Yan RT, Chen JS (1987) Butanol-Ethanol Dehydrogenase and Butanol-Ethanol-Iso-propanol Dehydrogenase: Different Alcohol Dehydrogenases in Two Strains of *Clostridium beijerinckii* (*Clostridium butylicum*). *Appl Environ Microbiol* 53: 697–703. PMID: [16347317](https://pubmed.ncbi.nlm.nih.gov/16347317/)

62. Youngleson JS, Jones WA, Jones DT, Woods DR (1989) Molecular analysis and nucleotide sequence of the *adh1* gene encoding an NADPH-dependent butanol dehydrogenase in the Gram-positive anaerobe *Clostridium acetobutylicum*. *Gene* 78: 355–364. PMID: [2673928](#)
63. Daniel R, Boenigk R, Gottschalk G (1995) Purification of 1,3-propanediol dehydrogenase from *Citrobacter freundii* and cloning, sequencing, and overexpression of the corresponding gene in *Escherichia coli*. *J Bacteriol* 177: 2151–2156. PMID: [7721705](#)
64. Elleuche S, Fodor K, Klippel B, von der HA, Wilmanns M, Antranikian G (2013) Structural and biochemical characterisation of a NAD(+)-dependent alcohol dehydrogenase from *Oenococcus oeni* as a new model molecule for industrial biotechnology applications. *Appl Microbiol Biotechnol* 97: 8963–8975. doi: [10.1007/s00253-013-4725-0](#) PMID: [23385476](#)
65. Luers F, Seyfried M, Daniel R, Gottschalk G (1997) Glycerol conversion to 1,3-propanediol by *Clostridium pasteurianum*: cloning and expression of the gene encoding 1,3-propanediol dehydrogenase. *FEMS Microbiol Lett* 154: 337–345. PMID: [9311132](#)
66. Marcal D, Rego AT, Carrondo MA, Enguita FJ (2009) 1,3-Propanediol dehydrogenase from *Klebsiella pneumoniae*: decameric quaternary structure and possible subunit cooperativity. *J Bacteriol* 191: 1143–1151. doi: [10.1128/JB.01077-08](#) PMID: [19011020](#)
67. Daubaras DL, Saido K, Chakrabarty AM (1996) Purification of hydroxyquinol 1,2-dioxygenase and maleylacetate reductase: the lower pathway of 2,4,5-trichlorophenoxyacetic acid metabolism by *Burkholderia cepacia* AC1100. *Appl Environ Microbiol* 62: 4276–4279. PMID: [8900023](#)
68. Perez-Pantoja D, Donoso RA, Sanchez MA, Gonzalez B (2009) Genuine genetic redundancy in maleylacetate-reductase-encoding genes involved in degradation of haloaromatic compounds by *Cupriavidus necator* JMP134. *Microbiology* 155: 3641–3651. doi: [10.1099/mic.0.032086-0](#) PMID: [19684066](#)
69. Seibert V, Stadler-Fritzsche K, Schlomann M (1993) Purification and characterization of maleylacetate reductase from *Alcaligenes eutrophus* JMP134(pJP4). *J Bacteriol* 175: 6745–6754. PMID: [8226615](#)
70. Seibert V, Kourbatova EM, Golovleva LA, Schlomann M (1998) Characterization of the maleylacetate reductase MacA of *Rhodococcus opacus* 1CP and evidence for the presence of an isofunctional enzyme. *J Bacteriol* 180: 3503–3508. PMID: [9657989](#)
71. Seibert V, Thiel M, Hinner IS, Schlomann M (2004) Characterization of a gene cluster encoding the maleylacetate reductase from *Ralstonia eutropha* 335T, an enzyme recruited for growth with 4-fluorobenzoate. *Microbiology* 150: 463–472. doi: [10.1099/mic.0.26602-0](#) PMID: [14766925](#)
72. Ma F, Wang T, Ma X, Wang P (2014) Identification and characterization of protein encoded by *orf382* as L-threonine dehydrogenase. *J Microbiol Biotechnol* 24: 748–755. PMID: [24651643](#)
73. Lyon RC, Johnston SM, Panopoulos A, Alzeer S, McGarvie G, Ellis EM (2009) Enzymes involved in the metabolism of gamma-hydroxybutyrate in SH-SY5Y cells: identification of an iron-dependent alcohol dehydrogenase ADHFe1. *Chem Biol Interact* 178: 283–287. doi: [10.1016/j.cbi.2008.10.025](#) PMID: [19013439](#)
74. Ward N, Moreno-Hagelsieb G (2014) Quickly finding orthologs as reciprocal best hits with BLAT, LAST, and UBLAST: how much do we miss? *PLoS One* 9: e101850. doi: [10.1371/journal.pone.0101850](#) PMID: [25013894](#)
75. Andersson JO, Hirt RP, Foster PG, Roger AJ (2006) Evolution of four gene families with patchy phylogenetic distributions: influx of genes into protist genomes. *BMC Evol Biol* 6: 27. doi: [10.1186/1471-2148-6-27](#) PMID: [16551352](#)
76. Greco M, Chiappetta A, Bruno L, Bitonti MB (2012) In *Posidonia oceanica* cadmium induces changes in DNA methylation and chromatin patterning. *J Exp Bot* 63: 695–709. doi: [10.1093/jxb/err313](#) PMID: [22058406](#)
77. Yang J, Zhang Y (2015) I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res* 43: W174–W181. doi: [10.1093/nar/gkv342](#) PMID: [25883148](#)
78. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25: 1605–1612. doi: [10.1002/jcc.20084](#) PMID: [15264254](#)
79. Bruchhaus I, Tannich E (1994) Purification and molecular characterization of the NAD(+)-dependent acetaldehyde/alcohol dehydrogenase from *Entamoeba histolytica*. *Biochem J* 303 (Pt 3): 743–748.
80. Pineda E, Encalada R, Olivos-Garcia A, Nequiz M, Moreno-Sanchez R, Saavedra E (2013) The bifunctional aldehyde-alcohol dehydrogenase controls ethanol and acetate production in *Entamoeba histolytica* under aerobic conditions. *FEBS Lett* 587: 178–184. doi: [10.1016/j.febslet.2012.11.020](#) PMID: [23201265](#)
81. Yang W, Li E, Kairong T, Stanley SL Jr. (1994) *Entamoeba histolytica* has an alcohol dehydrogenase homologous to the multifunctional *adhE* gene product of *Escherichia coli*. *Mol Biochem Parasitol* 64: 253–260. PMID: [7935603](#)

82. Kessler D, Leibrecht I, Knappe J (1991) Pyruvate-formate-lyase-deactivase and acetyl-CoA reductase activities of *Escherichia coli* reside on a polymeric protein particle encoded by *adhE*. *FEBS Lett* 281: 59–63. PMID: [2015910](#)
83. Kessler D, Herth W, Knappe J (1992) Ultrastructure and pyruvate formate-lyase radical quenching property of the multienzymic *AdhE* protein of *Escherichia coli*. *J Biol Chem* 267: 18073–18079. PMID: [1325457](#)
84. González-Segura L, Riveros-Rosas H, Julián-Sánchez A, Muñoz-Clares RA (2015) Residues that influence coenzyme preference in the aldehyde dehydrogenases. *Chem Biol Interact* 234: 59–74. doi: [10.1016/j.cbi.2014.12.039](#) PMID: [25601141](#)
85. Schwarzenbacher R, von DF, Canaves JM, Brinen LS, Dai X, Deacon AM, et al. (2004) Crystal structure of an iron-containing 1,3-propanediol dehydrogenase (TM0920) from *Thermotoga maritima* at 1.3 Å resolution. *Proteins* 54: 174–177. doi: [10.1002/prot.10594](#) PMID: [14705036](#)
86. Sulzenbacher G, Alvarez K, Van Den Heuvel RH, Versluis C, Spinelli S, Campanacci V, et al. (2004) Crystal structure of *E. coli* alcohol dehydrogenase YqhD: evidence of a covalently modified NADP coenzyme. *J Mol Biol* 342: 489–502. doi: [10.1016/j.jmb.2004.07.034](#) PMID: [15327949](#)
87. Antoine E, Rolland JL, Raffin JP, Dietrich J (1999) Cloning and over-expression in *Escherichia coli* of the gene encoding NADPH group III alcohol dehydrogenase from *Thermococcus hydrothermalis*. Characterization and comparison of the native and the recombinant enzymes. *Eur J Biochem* 264: 880–889. PMID: [10491136](#)
88. Li D, Stevenson KJ (1997) Purification and sequence analysis of a novel NADP(H)-dependent type III alcohol dehydrogenase from *Thermococcus* strain AN1. *J Bacteriol* 179: 4433–4437. PMID: [9209068](#)
89. Ying X, Grunden AM, Nie L, Adams MW, Ma K (2009) Molecular characterization of the recombinant iron-containing alcohol dehydrogenase from the hyperthermophilic Archaeon, *Thermococcus* strain ES1. *Extremophiles* 13: 299–311. doi: [10.1007/s00792-008-0217-z](#) PMID: [19115036](#)
90. Ma K, Adams MW (2001) Alcohol dehydrogenases from *Thermococcus litoralis* and *Thermococcus* strain ES-1. *Methods Enzymol* 331: 195–201. PMID: [11265461](#)
91. Ahvazi B, Coulombe R, Delarge M, Vedadi M, Zhang L, Meighen E, et al. (2000) Crystal structure of the NADP⁺-dependent aldehyde dehydrogenase from *Vibrio parvulus*: structural implications for cofactor specificity and affinity. *Biochem J* 349 Pt 3: 853–861.
92. Huang Y, Zhao KX, Shen XH, Chaudhry MT, Jiang CY, Liu SJ (2006) Genetic characterization of the resorcinol catabolic pathway in *Corynebacterium glutamicum*. *Appl Environ Microbiol* 72: 7238–7245. doi: [10.1128/AEM.01494-06](#) PMID: [16963551](#)
93. Perozich J, Kuo I, Wang BC, Boesch JS, Lindahl R, Hempel J (2000) Shifting the NAD/NADP preference in class 3 aldehyde dehydrogenase. *Eur J Biochem* 267: 6197–6203. PMID: [11012673](#)
94. Perozich J, Kuo I, Lindahl R, Hempel J (2001) Coenzyme specificity in aldehyde dehydrogenase. *Chem Biol Interact* 130–132: 115–124. PMID: [11306036](#)
95. Fujii T, Sato A, Okamoto Y, Yamauchi T, Kato S, Yoshida M, et al. (2016) The crystal structure of maleylacetate reductase from *Rhizobium* sp. strain MTP-10005 provides insights into the reaction mechanism of enzymes in its original family. *Proteins*.
96. Gaal AB, Neujahr HY (1980) Maleylacetate reductase from *Trichosporon cutaneum*. *Biochem J* 185: 783–786. PMID: [7387635](#)
97. Kaschabek SR, Reineke W (1993) Degradation of chloroaromatics: purification and characterization of maleylacetate reductase from *Pseudomonas* sp. strain B13. *J Bacteriol* 175: 6075–6081. PMID: [8407778](#)
98. Obradors N, Cabisco E, Aguilar J, Ros J (1998) Site-directed mutagenesis studies of the metal-binding center of the iron-dependent propanediol oxidoreductase from *Escherichia coli*. *Eur J Biochem* 258: 207–213. PMID: [9851711](#)
99. Elleuche S, Fodor K, von der HA, Klippel B, Wilmanns M, Antranikian G (2014) Group III alcohol dehydrogenase from *Pectobacterium atrosepticum*: insights into enzymatic activity and organization of the metal ion-containing region. *Appl Microbiol Biotechnol* 98: 4041–4051. doi: [10.1007/s00253-013-5374-z](#) PMID: [24265029](#)
100. Extance J, Crennell SJ, Eley K, Cripps R, Hough DW, Danson MJ (2013) Structure of a bifunctional alcohol dehydrogenase involved in bioethanol generation in *Geobacillus thermoglucosidarius*. *Acta Crystallogr D Biol Crystallogr* 69: 2104–2115. doi: [10.1107/S0907444913020349](#) PMID: [24100328](#)
101. Tamarit J, Cabisco E, Aguilar J, Ros J (1997) Differential inactivation of alcohol dehydrogenase isoenzymes in *Zymomonas mobilis* by oxygen. *J Bacteriol* 179: 1102–1104. PMID: [9023190](#)
102. Musille P, Ortlund E (2014) Structure of glycerol dehydrogenase from *Serratia*. *Acta Crystallogr F Struct Biol Commun* 70: 166–172. doi: [10.1107/S2053230X13034444](#) PMID: [24637749](#)

103. Ruzheinikov SN, Burke J, Sedelnikova S, Baker PJ, Taylor R, Bullough PA, et al. (2001) Glycerol dehydrogenase. structure, specificity, and mechanism of a family III polyol dehydrogenase. *Structure* 9: 789–802. PMID: [11566129](#)
104. Kardon T, Noel G, Vertommen D, Schaftingen EV (2006) Identification of the gene encoding hydroxyacid-oxoacid transhydrogenase, an enzyme that metabolizes 4-hydroxybutyrate. *FEBS Lett* 580: 2347–2350. doi: [10.1016/j.febslet.2006.02.082](#) PMID: [16616524](#)
105. Kaufman EE, Nelson T, Fales HM, Levin DM (1988) Isolation and characterization of a hydroxyacid-oxoacid transhydrogenase from rat kidney mitochondria. *J Biol Chem* 263: 16872–16879. PMID: [3182820](#)
106. Nelson T, Kaufman E, Kline J, Sokoloff L (1981) The extraneural distribution of gamma-hydroxybutyrate. *J Neurochem* 37: 1345–1348. PMID: [7299403](#)
107. Kaufman EE, Nelson T (1991) An overview of gamma-hydroxybutyrate catabolism: the role of the cytosolic NADP(+)-dependent oxidoreductase EC 1.1.1.19 and of a mitochondrial hydroxyacid-oxoacid transhydrogenase in the initial, rate-limiting step in this pathway. *Neurochem Res* 16: 965–974. PMID: [1784339](#)
108. Maitre M, Klein C, Mensah-Nyagan AG (2016) Mechanisms for the Specific Properties of gamma-Hydroxybutyrate in Brain. *Med Res Rev* 36: 363–388. doi: [10.1002/med.21382](#) PMID: [26739481](#)
109. Van CE, Plat L, Scharf MB, Leproult R, Cespedes S, L'Hermite-Baleriaux M, et al. (1997) Simultaneous stimulation of slow-wave sleep and growth hormone secretion by gamma-hydroxybutyrate in normal young Men. *J Clin Invest* 100: 745–753. doi: [10.1172/JCI119587](#) PMID: [9239423](#)
110. Koch JJ (2002) Performance-enhancing: substances and their use among adolescent athletes. *Pediatr Rev* 23: 310–317. PMID: [12205298](#)
111. McGivney BA, Eivers SS, MacHugh DE, MacLeod JN, O'Gorman GM, Park SD, et al. (2009) Transcriptional adaptations following exercise in thoroughbred horse skeletal muscle highlights molecular mechanisms that lead to muscle hypertrophy. *BMC Genomics* 10: 638. doi: [10.1186/1471-2164-10-638](#) PMID: [20042072](#)
112. Gu J, Orr N, Park SD, Katz LM, Sulimova G, MacHugh DE, et al. (2009) A genome scan for positive selection in thoroughbred horses. *PLoS One* 4: e5767. doi: [10.1371/journal.pone.0005767](#) PMID: [19503617](#)
113. Kim JY, Tillison KS, Zhou S, Lee JH, Smas CM (2007) Differentiation-dependent expression of *Adh1* in adipogenesis. *Arch Biochem Biophys* 464: 100–111. doi: [10.1016/j.abb.2007.04.018](#) PMID: [17559793](#)
114. Tae CH, Ryu KJ, Kim SH, Kim HC, Chun HK, Min BH, et al. (2013) Alcohol dehydrogenase, iron containing, 1 promoter hypermethylation associated with colorectal cancer differentiation. *BMC Cancer* 13: 142. doi: [10.1186/1471-2407-13-142](#) PMID: [23517143](#)
115. Guda C, Guda P, Fahy E, Subramaniam S (2004) MITOPRED: a web server for the prediction of mitochondrial proteins. *Nucleic Acids Res* 32: W372–W374. doi: [10.1093/nar/gkh374](#) PMID: [15215413](#)
116. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, et al. (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res* 35: W585–W587. doi: [10.1093/nar/gkm259](#) PMID: [17517783](#)
117. Petsalaki EI, Bagos PG, Litou ZI, Hamodrakas SJ (2006) PredSL: a tool for the N-terminal sequence-based prediction of protein subcellular localization. *Genomics Proteomics Bioinformatics* 4: 48–55. doi: [10.1016/S1672-0229\(06\)60016-8](#) PMID: [16689702](#)
118. Kardon T, Noel G, Vertommen D, Schaftingen EV (2006) Identification of the gene encoding hydroxyacid-oxoacid transhydrogenase, an enzyme that metabolizes 4-hydroxybutyrate. *FEBS Lett* 580: 2347–2350. doi: [10.1016/j.febslet.2006.02.082](#) PMID: [16616524](#)
119. Sakurai M, Tohda H, Kumagai H, Giga-Hama Y (2004) A distinct type of alcohol dehydrogenase, *adh4+*, complements ethanol fermentation in an *adh1*-deficient strain of *Schizosaccharomyces pombe*. *FEMS Yeast Res* 4: 649–654. doi: [10.1016/j.femsyr.2003.12.009](#) PMID: [15040954](#)
120. de Smidt O, du Preez JC, Albertyn J (2012) Molecular and physiological aspects of alcohol dehydrogenases in the ethanol metabolism of *Saccharomyces cerevisiae*. *FEMS Yeast Res* 12: 33–47. doi: [10.1111/j.1567-1364.2011.00760.x](#) PMID: [22094012](#)
121. Drewke C, Ciriacy M (1988) Overexpression, purification and properties of alcohol dehydrogenase IV from *Saccharomyces cerevisiae*. *Biochim Biophys Acta* 950: 54–60. PMID: [3282541](#)
122. Ganzhorn AJ, Green DW, Hershey AD, Gould RM, Plapp BV (1987) Kinetic characterization of yeast alcohol dehydrogenases. Amino acid residue 294 and substrate specificity. *J Biol Chem* 262: 3754–3761. PMID: [3546317](#)
123. Espinosa A, Yan L, Zhang Z, Foster L, Clark D, Li E, et al. (2001) The bifunctional *Entamoeba histolytica* alcohol dehydrogenase 2 (*EhADH2*) protein is necessary for amebic growth and survival and

- requires an intact C-terminal domain for both alcohol dehydrogenase and acetaldehyde dehydrogenase activity. *J Biol Chem* 276: 20136–20143. doi: [10.1074/jbc.M101349200](https://doi.org/10.1074/jbc.M101349200) PMID: [11274185](https://pubmed.ncbi.nlm.nih.gov/11274185/)
124. Dan M, Wang CC (2000) Role of alcohol dehydrogenase E (ADHE) in the energy metabolism of *Giardia lamblia*. *Mol Biochem Parasitol* 109: 25–36. PMID: [10924754](https://pubmed.ncbi.nlm.nih.gov/10924754/)
 125. Atteia A, van Lis R, Mendoza-Hernandez G, Henze K, Martin W, Riveros-Rosas H, et al. (2003) Bifunctional aldehyde/alcohol dehydrogenase (ADHE) in chlorophyte algal mitochondria. *Plant Mol Biol* 53: 175–188. doi: [10.1023/B:PLAN.0000009274.19340.36](https://doi.org/10.1023/B:PLAN.0000009274.19340.36) PMID: [14756315](https://pubmed.ncbi.nlm.nih.gov/14756315/)
 126. Shasmal M, Dey S, Shaikh TR, Bhakta S, Sengupta J (2016) E. coli metabolic protein aldehyde-alcohol dehydrogenase-E binds to the ribosome: a unique moonlighting action revealed. *Sci Rep* 6: 19936. doi: [10.1038/srep19936](https://doi.org/10.1038/srep19936) PMID: [26822933](https://pubmed.ncbi.nlm.nih.gov/26822933/)
 127. Buswell JA, Eriksson K-E (1979) Aromatic ring cleavage by the white-rot fungus *Sporotrichum pulverulentum*. *FEBS Lett* 104: 258–260.
 128. Jones KH, Trudgill PW, Hopper DJ (1995) Evidence of two pathways for the metabolism of phenol by *Aspergillus fumigatus*. *Arch Microbiol* 163: 176–181. PMID: [7778974](https://pubmed.ncbi.nlm.nih.gov/7778974/)
 129. Patel TR, Hameed N, Armstrong S (1992) Metabolism of gallate in *Penicillium simplicissimum*. *J Basic Microbiol* 32: 233–240. PMID: [1460567](https://pubmed.ncbi.nlm.nih.gov/1460567/)
 130. Rieble S, Joshi DK, Gold MH (1994) Purification and characterization of a 1,2,4-trihydroxybenzene 1,2-dioxygenase from the basidiomycete *Phanerochaete chrysosporium*. *J Bacteriol* 176: 4838–4844. PMID: [8050996](https://pubmed.ncbi.nlm.nih.gov/8050996/)
 131. Shailubhai K, Somayaji R, Rao NN, Modi VV (1983) Metabolism of resorcinol and salicylate in *Aspergillus niger*. *Experientia* 39: 70–72. PMID: [6825782](https://pubmed.ncbi.nlm.nih.gov/6825782/)
 132. Sparnins VL, Burbee DG, Dagley S (1979) Catabolism of L-tyrosine in *Trichosporon cutaneum*. *J Bacteriol* 138: 425–430. PMID: [571434](https://pubmed.ncbi.nlm.nih.gov/571434/)
 133. Seo JA, Proctor RH, Plattner RD (2001) Characterization of four clustered and coregulated genes associated with fumonisin biosynthesis in *Fusarium verticillioides*. *Fungal Genet Biol* 34: 155–165. doi: [10.1006/fgbi.2001.1299](https://doi.org/10.1006/fgbi.2001.1299) PMID: [11728154](https://pubmed.ncbi.nlm.nih.gov/11728154/)
 134. Li Y, Lou L, Cerny RL, Butchko RA, Proctor RH, Shen Y, et al. (2013) Tricarballic ester formation during biosynthesis of fumonisin mycotoxins in. *Mycology* 4: 179–186. doi: [10.1080/21501203.2013.874540](https://doi.org/10.1080/21501203.2013.874540) PMID: [24587959](https://pubmed.ncbi.nlm.nih.gov/24587959/)
 135. Butchko RA, Plattner RD, Proctor RH (2006) Deletion analysis of FUM genes involved in tricarballic ester formation during fumonisin biosynthesis. *J Agric Food Chem* 54: 9398–9404. doi: [10.1021/jf0617869](https://doi.org/10.1021/jf0617869) PMID: [17147424](https://pubmed.ncbi.nlm.nih.gov/17147424/)
 136. Pereto J, Lopez-Garcia P, Moreira D (2004) Ancestral lipid biosynthesis and early membrane evolution. *Trends Biochem Sci* 29: 469–477. doi: [10.1016/j.tibs.2004.07.002](https://doi.org/10.1016/j.tibs.2004.07.002) PMID: [15337120](https://pubmed.ncbi.nlm.nih.gov/15337120/)
 137. Moran Y, Fredman D, Szczesny P, Grynberg M, Technau U (2012) Recurrent horizontal transfer of bacterial toxin genes to eukaryotes. *Mol Biol Evol* 29: 2223–2230. doi: [10.1093/molbev/mss089](https://doi.org/10.1093/molbev/mss089) PMID: [22411854](https://pubmed.ncbi.nlm.nih.gov/22411854/)
 138. Martin-Duran JM, de Mendoza A, Sebe-Pedros A, Ruiz-Trillo I, Hejnal A (2013) A broad genomic survey reveals multiple origins and frequent losses in the evolution of respiratory hemerythrins and hemo-cyanins. *Genome Biol Evol* 5: 1435–1442. doi: [10.1093/gbe/evt102](https://doi.org/10.1093/gbe/evt102) PMID: [23843190](https://pubmed.ncbi.nlm.nih.gov/23843190/)
 139. Yang S, Bourne PE (2009) The evolutionary history of protein domains viewed by species phylogeny. *PLoS One* 4: e8378. doi: [10.1371/journal.pone.0008378](https://doi.org/10.1371/journal.pone.0008378) PMID: [20041107](https://pubmed.ncbi.nlm.nih.gov/20041107/)
 140. Elias M, Novotny M (2008) cpRAS: a novel circularly permuted RAS-like GTPase domain with a highly scattered phylogenetic distribution. *Biol Direct* 3: 21. doi: [10.1186/1745-6150-3-21](https://doi.org/10.1186/1745-6150-3-21) PMID: [18510733](https://pubmed.ncbi.nlm.nih.gov/18510733/)
 141. Chan CX, Bhattacharya D, Reyes-Prieto A (2012) Endosymbiotic and horizontal gene transfer in microbial eukaryotes: Impacts on cell evolution and the tree of life. *Mob Genet Elements* 2: 101–105. doi: [10.4161/mge.20110](https://doi.org/10.4161/mge.20110) PMID: [22934244](https://pubmed.ncbi.nlm.nih.gov/22934244/)
 142. Koonin EV (2010) The origin and early evolution of eukaryotes in the light of phylogenomics. *Genome Biol* 11: 209. doi: [10.1186/gb-2010-11-5-209](https://doi.org/10.1186/gb-2010-11-5-209) PMID: [20441612](https://pubmed.ncbi.nlm.nih.gov/20441612/)
 143. Shapiro JA (2016) Nothing in Evolution Makes Sense Except in the Light of Genomics: Read-Write Genome Evolution as an Active Biological Process. *Biology (Basel)* 5.
 144. Armengaud J, Timmis KN, Wittich RM (1999) A functional 4-hydroxysalicylate/hydroxyquinol degradative pathway gene cluster is linked to the initial dibenzo-p-dioxin pathway genes in *Sphingomonas* sp. strain RW1. *J Bacteriol* 181: 3452–3461. PMID: [10348858](https://pubmed.ncbi.nlm.nih.gov/10348858/)
 145. Daubaras DL, Hershberger CD, Kitano K, Chakrabarty AM (1995) Sequence analysis of a gene cluster involved in metabolism of 2,4,5-trichlorophenoxyacetic acid by *Burkholderia cepacia* AC1100. *Appl Environ Microbiol* 61: 1279–1289. PMID: [7538273](https://pubmed.ncbi.nlm.nih.gov/7538273/)

146. Endo R, Kamakura M, Miyauchi K, Fukuda M, Ohtsubo Y, Tsuda M, et al. (2005) Identification and characterization of genes involved in the downstream degradation pathway of gamma-hexachlorocyclohexane in *Sphingomonas paucimobilis* UT26. *J Bacteriol* 187: 847–853. doi: [10.1128/JB.187.3.847-853.2005](https://doi.org/10.1128/JB.187.3.847-853.2005) PMID: [15659662](https://pubmed.ncbi.nlm.nih.gov/15659662/)
147. Camara B, Nikodem P, Bielecki P, Bobadilla R, Junca H, Pieper DH (2009) Characterization of a gene cluster involved in 4-chlorocatechol degradation by *Pseudomonas reinekei* MT1. *J Bacteriol* 191: 4905–4915. doi: [10.1128/JB.00331-09](https://doi.org/10.1128/JB.00331-09) PMID: [19465655](https://pubmed.ncbi.nlm.nih.gov/19465655/)
148. Yoshida M, Oikawa T, Obata H, Abe K, Mihara H, Esaki N (2007) Biochemical and genetic analysis of the gamma-resorcyate (2,6-dihydroxybenzoate) catabolic pathway in *Rhizobium* sp. strain MTP-10005: identification and functional analysis of its gene cluster. *J Bacteriol* 189: 1573–1581. doi: [10.1128/JB.01675-06](https://doi.org/10.1128/JB.01675-06) PMID: [17158677](https://pubmed.ncbi.nlm.nih.gov/17158677/)
149. Youngleson JS, Santangelo JD, Jones DT, Woods DR (1988) Cloning and Expression of a Clostridium acetobutylicum Alcohol Dehydrogenase Gene in *Escherichia coli*. *Appl Environ Microbiol* 54: 676–682. PMID: [16347579](https://pubmed.ncbi.nlm.nih.gov/16347579/)
150. Blodgett JA, Zhang JK, Metcalf WW (2005) Molecular cloning, sequence analysis, and heterologous expression of the phosphinothricin tripeptide biosynthetic gene cluster from *Streptomyces viridochromogenes* DSM 40736. *Antimicrob Agents Chemother* 49: 230–240. doi: [10.1128/AAC.49.1.230-240.2005](https://doi.org/10.1128/AAC.49.1.230-240.2005) PMID: [15616300](https://pubmed.ncbi.nlm.nih.gov/15616300/)
151. Blodgett JA, Thomas PM, Li G, Velasquez JE, van der Donk WA, Kelleher NL, et al. (2007) Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide. *Nat Chem Biol* 3: 480–485. doi: [10.1038/nchembio.2007.9](https://doi.org/10.1038/nchembio.2007.9) PMID: [17632514](https://pubmed.ncbi.nlm.nih.gov/17632514/)
152. Woodyer RD, Shao Z, Thomas PM, Kelleher NL, Blodgett JA, Metcalf WW, et al. (2006) Heterologous production of fosfomycin and identification of the minimal biosynthetic gene cluster. *Chem Biol* 13: 1171–1182. doi: [10.1016/j.chembiol.2006.09.007](https://doi.org/10.1016/j.chembiol.2006.09.007) PMID: [17113999](https://pubmed.ncbi.nlm.nih.gov/17113999/)
153. Woodyer RD, Li G, Zhao H, van der Donk WA (2007) New insight into the mechanism of methyl transfer during the biosynthesis of fosfomycin. *Chem Commun (Camb)* 359–361.
154. Struys EA, Verhoeven NM, Ten Brink HJ, Wickenhagen WV, Gibson KM, Jakobs C (2005) Kinetic characterization of human hydroxyacid-oxoacid transhydrogenase: relevance to D-2-hydroxyglutaric and gamma-hydroxybutyric acidurias. *J Inher Metab Dis* 28: 921–930. doi: [10.1007/s10545-005-0114-x](https://doi.org/10.1007/s10545-005-0114-x) PMID: [16435184](https://pubmed.ncbi.nlm.nih.gov/16435184/)
155. Iwaki H, Hasegawa Y, Wang S, Kayser MM, Lau PC (2002) Cloning and characterization of a gene cluster involved in cyclopentanol metabolism in *Comamonas* sp. strain NCIMB 9872 and biotransformations effected by *Escherichia coli*-expressed cyclopentanone 1,2-monooxygenase. *Appl Environ Microbiol* 68: 5671–5684. doi: [10.1128/AEM.68.11.5671-5684.2002](https://doi.org/10.1128/AEM.68.11.5671-5684.2002) PMID: [12406764](https://pubmed.ncbi.nlm.nih.gov/12406764/)
156. Gattis SG, Chung HS, Trent MS, Raetz CR (2013) The origin of 8-amino-3,8-dideoxy-D-manno-octulosonic acid (Kdo8N) in the lipopolysaccharide of *Shewanella oneidensis*. *J Biol Chem* 288: 9216–9225. doi: [10.1074/jbc.M113.453324](https://doi.org/10.1074/jbc.M113.453324) PMID: [23413030](https://pubmed.ncbi.nlm.nih.gov/23413030/)
157. Brzostowicz PC, Blasko MS, Rouviere PE (2002) Identification of two gene clusters involved in cyclohexanone oxidation in *Brevibacterium epidermidis* strain HCU. *Appl Microbiol Biotechnol* 58: 781–789. doi: [10.1007/s00253-002-0968-x](https://doi.org/10.1007/s00253-002-0968-x) PMID: [12021799](https://pubmed.ncbi.nlm.nih.gov/12021799/)