A polyextremophilic alcohol dehydrogenase from the Atlantis II Deep Red Sea brine pool

Anastassja L. Akal[1,2], Ram Karan[1], Adrian Hohl[1,2], Intikhab Alam[3], Malvina Vogler[1,2], Stefan W. Grötzinger[3,4], Jörg Eppinger[1] and Magnus Rueping[1,5]

1 KAUST Catalysis Center (KCC), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
2 Center for Integrated Protein Science Munich at the Department of Chemistry, Technical University of Munich (TUM), Garching, Germany
3 Computational Bioscience Research Center (CBRC), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
4 Institute of Biochemical Engineering, Technical University of Munich (TUM), Garching, Germany
5 Institute of Organic Chemistry, RWTH Aachen, Aachen, Germany

Enzymes originating from hostile environments offer exceptional stability under industrial conditions and are therefore highly in demand. Using single-cell genome data, we identified the alcohol dehydrogenase (ADH) gene, *adh/a1a*, from the Atlantis II Deep Red Sea brine pool. ADH/A1a is highly active at elevated temperatures and high salt concentrations (optima at 70 °C and 4 M KCl) and withstands organic solvents. The polyextremophilic ADH/A1a exhibits a broad substrate scope including aliphatic and aromatic alcohols and is able to reduce cinnamyl-methyl-ketone and raspberry ketone in the reverse reaction, making it a possible candidate for the production of chiral compounds. Here, we report the affiliation of ADH/A1a to a rare enzyme family of microbial cinnamyl alcohol dehydrogenases and explain unique structural features for halo- and thermoadaptation.

Alcohol dehydrogenases (ADHs, EC 1.1.1.1) are enzymes that catalyze the reversible dehydrogenation of alcohols to aldehydes or ketones by consuming the cofactor nicotinamide adenine dinucleotide or its phosphate ($NAD^+/NADP^+$). In the biochemical industry, the natural regio- and enantioselectivity of ADHs are exploited to produce chiral compounds for pharmaceuticals and fine chemicals and are therefore worthwhile studying [1,2]. One of the major drawbacks of conventional ADHs is their poor performance under industrial conditions such as extreme temperatures, pH, and the presence of salts and organic solvents, as they tend to lose their activity and stability [3]. However, these operational conditions are a prerequisite to increase the solubility of hydrophobic compounds, reduce viscosity and accelerate the reaction, or influence the enantioselectivity [4,5]. Besides the catalytic reaction, the enzymes have to endure unnatural conditions during transportation, storage, and formulation [2]. Therefore, industry demands the development of more robust and stereoselective ADHs.

In addition to protein-engineering methods, the exploration of enzymes derived from extremophilic organisms is a major pool for highly stable ADHs [6]. Extremophiles are naturally adapted to survive and grow in hostile environments, and their evolved enzymes, also called extremozymes, are not only stable but highly active under extreme conditions [7]. The structurally promoted resistance of extremozymes is still poorly understood and currently under investigation [8].

Abbreviations

ADH, alcohol dehydrogenase; CAD, cinnamyl alcohol dehydrogenase; CDD, conserved domain database; MDR, medium-chain dehydrogenase/reductase; SAG, single amplified genome.

Extremozymes derived from marine microorganisms are reported to be stable under harsh conditions including high ionic strengths and solvents and further show novel functionalities [9]. One of the most extreme marine environments is the Red Sea brine pools, characterized by high salinity, elevated temperatures, high concentrations of heavy metals, hydrostatic pressure, and lack of oxygen and light [10]. Several enzymes derived from brine pools exhibit polyextremophilic characteristics and are therefore worth to be studied to elucidate their structure–stability relationship and to find potential enzymes for industrial applications [11,12].

For the discovery of novel extremozymes, we explored the largest Red Sea brine pool, the Atlantis II Deep, which exhibits temperatures up to 68 °C and salinities up to 26%. Several genes were annotated from single amplified genomes (SAGs) of uncultured single cells derived from water samples and are open for exploration [13]. Here, we characterized the ADH gene, *adh/a1a*, from an uncharacterized *MSBL1 archaeon* collected from the Atlantis II Deep interphase (2036 m, 16.8% salt, 63 °C, pH 5.3). ADH/A1a was overexpressed in the halophilic *Haloferax volcanii* strain under high-salt condition, resembling its natural environment. The purified ADH/A1 was highly active in high salt concentrations and elevated temperatures and resistant toward organic solvent additives. Furthermore, ADH/A1a revealed to have a broad substrate spectrum, with a preference for primary long-chain and aromatic alcohols such as cinnamyl alcohol and its derivatives. The *in silico* structural analysis of ADH/A1a and homologs uncovered the evolved structural features of its halo-thermophilic adaptation and may lead to engineering strategies for robust and rational designed biocatalysts. The phylogenetic analysis revealed that ADH/A1a shares features of classical cinnamyl alcohol dehydrogenases (CADs) found in plants and propanol-preferring ADHs [14,15].

Materials and methods

Chemicals, strains, and vectors

The restriction endonucleases and chemically competent *Escherichia coli* cells were purchased from New England Biolabs (Ipswich, MA, USA). Most other chemicals were purchased from Sigma (St. Louis, MO, USA). *Haloferax volcanii* H1895 and its corresponding vector pTA963 were kindly provided by Thorsten Allers (Institute of Genetics, University of Nottingham, Queen's Medical Centre, UK) and Eva Strillinger (Institute of Biochemical Engineering, TUM, Germany) [16,17].

Source and annotation of SAGs from the Red Sea brine pools

Samples were collected from different depths of several Red Sea brine pools. From these samples, single cells were sorted and lysed, and the whole genome was amplified and sequenced [18]. The SAGs were annotated and the putative genes were fed into the INDIGO database [13]. The gene *adh/a1* was derived from the SAG of an unclassified *MSBL1* (Mediterranean Sea Brine Lake 1) archaeon of the Atlantis II Deep interphase and identified by the profile pattern match algorithm [19].

Construction of expression plasmid pTA963-*adh/a1a*

The *adh/a1* gene was codon-optimized using the Java codon adaptation online tool (JCat) for the codon usage of *Halobacterium* sp. NRC-1, which is similar to the codon usage of *H. volcanii* [20,21]. The gene was synthesized by Geneart (Regensburg, Germany) and cloned into the plasmid pTA963. The shortened gene *adh/a1a*, which excluded the N-terminal nonsense sequence of *adh/a1*, was created using the Gibson assembly cloning kit from NEB (Ipswich, MA, USA). The cloned construct pTA963-*adh/a1a* was transformed into the biofilm-free expression strain *H. volcanii* H1895 using the protocol from the Halohandbook [22]. The shortened gene *adh/a1a* was deposited in National Center for Biotechnology Informations (NCBI's) GenBank (GenBank: KXB02677) [18].

Production and purification of ADH/A1a

Protein production in *H. volcanii* and purification using a HisTrap HP column (GE Healthcare, Little Chalfont, UK) were performed as described with some buffer modifications [17]. The running and storage buffer contained additionally 10% (v/v) of glycerol (running/storage buffer: 10 mM HEPES/NaOH, pH 7.5, 2 M NaCl, 10% v/v glycerol). After purification, the active fractions were concentrated ($\sim$ 2–3 mg·mL^{-1}) using an Amicon Ultracel® 30 000 MWCO (Millipore, Billerica, MA, USA).

For further purification, the concentrated protein solution was loaded onto a HiLoad 16/60 Superdex 75 prep grade preparative column (GE Healthcare) with a flow rate of 1 mL·min^{-1}. The active protein fractions were analyzed by sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS/PAGE), and the purity of the protein was calculated using IMAGEJ version 1.51 (Wayne Rasband; NIH, Bethesda, MD, USA). The protein concentration was measured using a NanoDrop 2000c spectrophotometer (Thermo Scientific, Waltham, MA, USA). The extinction coefficient (26 930 M^{-1}·cm^{-1}) and molecular weight (37.6 kDa) of ADH/A1a were calculated using ProtParam [23].

Tryptic digest and LC-MS/MS analysis of ADH/A1a

The identification of ADH/A1a was performed by LC-MS/MS analysis, using a Coomassie-stained protein band, excised from a SDS/PAGE, destained, and digested using the In-Gel Tryptic Digest Kit (Thermo Fisher Scientific, Waltham, MA, USA). Peptides were measured using a LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific) and analyzed using MASCOT v2.3 (Matrix Sciences Ltd, Manchester, UK).

Determination of ADH/A1a activity

The activity of ADH/A1a was determined by measuring the absorbance of NAD(P)H at 340 nm with a spectrophotometer (Cary 60; Agilent, Santa Clara, CA, USA). The reaction was performed in 1-mL semi-micro PMMA cuvettes (Sigma-Aldrich) at 60 °C for 30 min. Unless otherwise specified, the oxidation reaction was performed in 3 M KCl, 50 mM glycine/NaOH buffer (pH 10.0) with 10 mM NAD(P)$^{+}$, and 0.2% (v/v) substrate; the reduction reaction was performed in 3 M KCl, 50 mM K$_2$HPO$_4$/citric acid buffer (pH 6.0) with 0.2 mM NAD(P)H, and 0.4% (v/v) substrate. Before the reaction, the ADH/A1a (0.4–1 μM) was incubated with zinc sulfate (0.4 mM) in the reaction mixture at room temperature for 15 min. The reaction was induced by the addition of the substrate. The kinetic parameters were calculated with GRAPHPAD PRISM v7 (GraphPad Software Inc., La Jolla, CA, USA).

Thermal stability of ADH/A1a

The thermal melting curve was determined using differential static light scattering measured with a Stargazer-2 (Harbinger Biotechnology, Toronto, ON, Canada) with a 1.5 mg·mL^{-1} ADH/A1a sample in dialysis buffer (10 mM HEPES/NaOH, pH 7.5, 2 M NaCl). The enzyme was pre-incubated with zinc sulfate for 30 min and dialyzed against dialysis buffer for 2 h at RT before measurement. The melting temperature T_m was calculated using GRAPHPAD PRISM (GraphPad Software Inc.).

Effect of various parameters on the ADH/A1a activity

The effect of various metal ions was tested with metal-free ADH/A1a. To remove nonspecific bound metal ions, the protein was dialyzed overnight at 4 °C against dialysis buffer containing 10 mM EDTA. Afterward, EDTA was removed by further dialysis against EDTA-free dialysis buffer. ADH/A1a was incubated with the metal ions for 30 min at room temperature before reaction.

Different NAD(P)$^{+}$ concentrations were tested in the oxidation reaction. The reaction was performed under standard conditions (60 °C, 3 M KCl, pH 10) using cinnamyl alcohol as substrate.

Zinc content of active ADH/A1a solution

An ADH/A1a solution (c = 1.7 mg·mL^{-1}, V = 1 mL) was incubated with 20 μL of 100 mM ZnSO$_4$ solution for 1 h at room temperature and subsequently dialyzed against dialysis buffer (3 × buffer exchange) at 4 °C for 24 h. The samples were recorded with a 5100 ICP-OES instrument (Agilent) coupled to an SPS 4 Autosampler (Agilent), using argon as gas supply. The digestion was done with 8 mL of nitric acid at 240 °C and 100 bar using an UltraWAVE apparatus (Milestone, Shelton, CT, USA). Prior to digestion of the samples, vessel cleaning was performed using only nitric acid (5 mL). All the samples and measurements were taken in duplicates.

Bioinformatic analysis

The analysis of the amino acid sequence and the classification of the protein family and subfamily were done using the CD-search tool of the Conserved Domain Database (CDD) of the NCBI [14,24]. The subdomains were identified using INTERPROSCAN [25]. The secondary structure elements of ADH/A1a were predicted using JPRED 4 [26]. Sequence-based homologous ADHs with resolved crystal structures were identified using SWISS-MODEL and I-TASSER [27,28]. The homology model was created using SWISS-MODEL and was evaluated by the GMQE and QMEAN scores [29]. The QMEANDisCo method was used to validate the local quality of the model [30]. The electrostatic surface potential was calculated using APBS of the PDB2PQR 2.0 with the default values and applied to the structures using the APBS plugin of PYMOL v1.7 (Schrödinger LLC, New York, NY, USA) [31,32]. The surface-accessible residues were calculated with the Swiss-PDB Viewer 4.1.1 using ≥ 30% surface accessibility [33].

For the phylogenetic analysis, experimentally confirmed ADH sequences of various subfamilies were selected using UniProt [34]. Uncharacterized sequences of the CAD2 were selected from the CDD [24,35]. Sequence alignments were done using MAFFT version 7 [34]. The phylogenetic tree was constructed with FastTree2 using the maximum-likelihood method and the Jones–Taylor–Thornton substitution model (1000 bootstraps, default settings) [36]. Functional annotation and possible metabolic pathways were investigated using BlastKOALA of the KEGG [37].

Results

Analysis of *adh/a1a*

The gene *adh/a1* was annotated from a single amplified genome (SAG) of an uncultured *MSBL1* archaeon, which was derived from the Atlantis II Deep Red Sea brine pool [13]. Our analysis of the gene sequence revealed the rare alternative start codon UUG, which

was previously missed by the annotation algorithm and is known to be used by haloarchaea [38]. Therefore, we deleted the N-terminal 20-amino acid-long random coil region of *adh/a1* and named the shortened gene *adh/a1a*. We classified ADH/A1a as a member of the zinc-binding medium-chain alcohol dehydrogenase (MDR) superfamily, exhibiting highly conserved domains of the CAD family [15]. Based on the conserved features of the CAD family, the catalytic and structural zinc-binding sites, as well as putative NAD(P)H- and substrate-binding sites, were predicted. Furthermore, the N-terminal GroES-like domain (position 1–152) and the C-terminal NAD(P)-binding domain (position 153–294) were identified (Fig. S1).

Production and purification of ADH/A1a

Initial attempts to produce ADH/A1a in *E. coli* under low-salt conditions resulted in an insoluble and inactive protein aggregate (data not shown). Therefore, ADH/A1a was expressed in the halophilic *H. volcanii* H1895 expression system with an N-terminal his$_6$-tag [17]. The expression and purification in the presence of high salt concentrations resulted in soluble and active ADH/A1a ($\sim$ 15 mg L^{-1} yield; Figs S2 and S3A). In contrast, purification at a low salt concentration (≤ 0.5 M NaCl) resulted in inactivation of the enzyme; however, it increased the purity (Fig. S3B). The native PAGE analysis revealed the oligomeric state of the enzyme (Fig. S4). The identity of the ADH/A1a and residual proteins was confirmed by tryptic digest and LC-MS/MS analysis (Fig. S5, Table S1).

Characterization of ADH/A1a

The deletion of the N-terminal nonsense region of the ADH/A1 did not affect the activity or stability of the ADH (data not shown). Therefore, the correct annotated ADH/A1a was used for characterization.

The activity of the purified ADH/A1a was assayed over a broad range of salt (NaCl and KCl) concentrations, temperatures, and pH values to determine the optimal reaction conditions. ADH/A1a was active in a salt concentration range of 1.5–5 M, with the maximum activity found at 4 M salt; higher activities were observed in KCl than in NaCl (Fig. 1A). The effect of temperature on the enzyme activity was assayed from 10 to 80 °C, with activity peaking at the relatively high temperature of 70 °C (Fig. 1B). Substantial activity was observed at 80 °C; however, over time protein aggregation occurred that is in agreement with the determined T_m of $\sim$ 73 °C (Fig. S6). The enzyme showed very divergent pH activity optima in the

oxidation and reduction reactions, at pH 10.0 and pH 6.5, respectively (Fig. 1C,D). A thermodynamic analysis of the reactions revealed concordantly that the oxidation reaction is energetically favored at a basic pH, whereas the reduction reaction is favored at an acidic pH (Fig. S7; calculated with eQuilibrator [39]).

The catalytically active metal was determined by incubating ADH/A1a with different metal ions before catalysis. The enzyme activity increased 3.5-fold with the addition of zinc(II) ions. In contrast, manganese(II) had no effect, and nickel(II), magnesium(II), and cobalt(II) had adverse effects on the enzyme activity. The addition of either copper(II) or chromium(II) resulted in a total loss of enzyme activity (Fig. 1E). Metal-free ADH/A1a showed activity exclusively with zinc(II), confirming its zinc dependency (Fig. S8A). Relatively high zinc concentrations (≥ 0.1 mM) were required in the reaction buffer to reach the optimal activity of the enzyme (Fig. S8B). Furthermore, a high content of bound zinc ions (18.9 g-atoms per ADH/A1a subunit) was determined in the ADH/A1a solution by ICP-OES (Table S2).

The cofactors NAD$^+$/NADH and NADP$^+$/NADPH were almost equally accepted (Fig. 1F). The optimal cofactor concentration for the oxidation reaction was 10 mM, showing substrate inhibition with higher concentrations (Fig. S8C).

Alcohol dehydrogenase/A1a oxidized a broad range of substrates including small-, medium-, and long-chain alcohols with a preference for long primary alcohols (Table 1A). Notably, several phenol and terpene alcohols were accepted as substrate. The highest reaction rates were achieved with cinnamyl alcohol ($\sim$ 0.49 µmol·min^{-1}·mg^{-1}), followed by 1,5-pentanediol and 1-heptanol ($\sim$ 0.29 µmol·min^{-1}·mg^{-1} for both). Remarkably, the secondary alcohol 3-buten-2-ol ($\sim$ 0.05 µmol·min^{-1}·mg^{-1}) was also oxidized, though at slow rates. In the reduction reaction, a conversion was exclusively found for cinnamaldehyde (0.31 µmol·min^{-1}·mg^{-1}), raspberry ketone ($\sim$0.25 µmol·min^{-1}·mg^{-1}), and cinnamyl-methyl-ketone ($\sim$0.40 µmol·min^{-1}·mg^{-1}; Table 1B).

Alcohol dehydrogenase/A1a followed Michaelis–Menten kinetics in the oxidation reaction and showed a first-order reaction (Fig. S9). The kinetic parameters confirmed the preference for cinnamyl alcohol as a substrate (Table 2).

Effect of salt and organic solvents on ADH/A1a stability

Alcohol dehydrogenase/A1a was remarkably stable at high salt concentrations (0.8–3 M NaCl); however, its stability significantly decreased at low salt concentrations
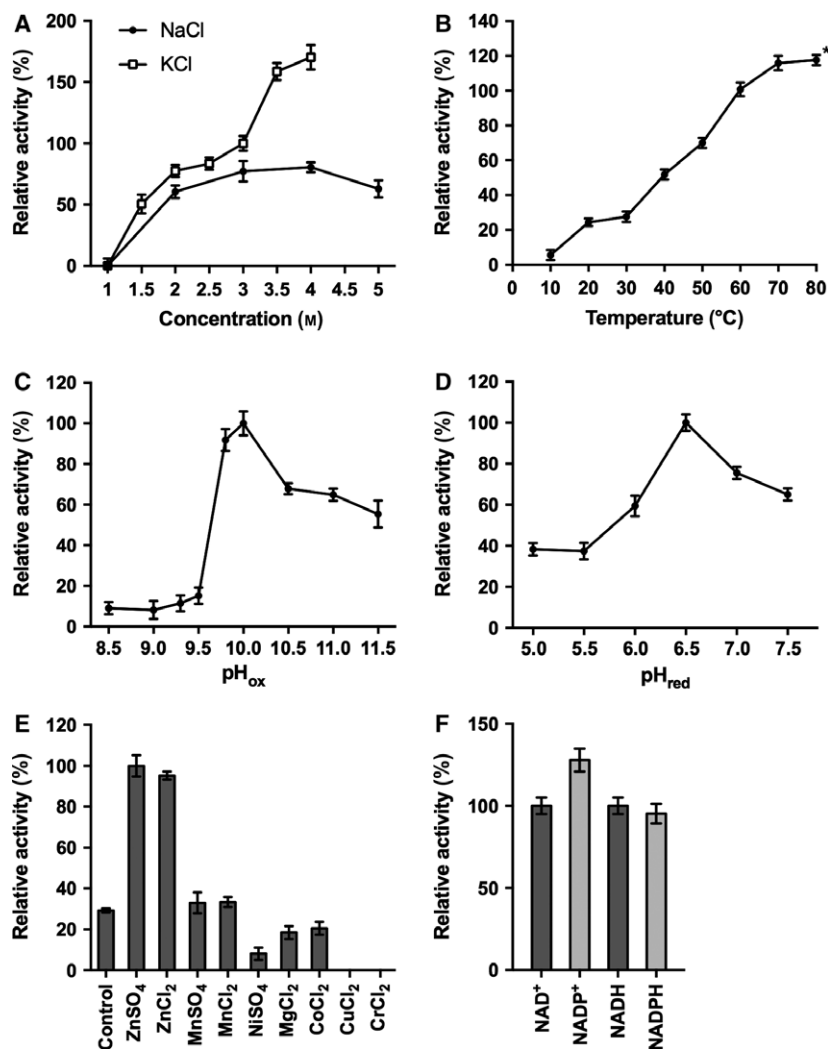
Fig. 1. Effect of various factors on the enzyme activity. (A) Salt concentration (saturation limit: 4 M for KCl, 5 M for NaCl), (B) temperature, (C) pH in oxidation, (D) pH in reduction reaction, and (E) metal ions (0.2 mM), protein without additive as control; and (F) cofactors for oxidation/reduction (oxidation: 100% = 0.49 $\mu mol \cdot min^{-1} \cdot mg^{-1}$; reduction: 100% = 0.31 $\mu mol \cdot min^{-1} \cdot mg^{-1}$). Error bars indicate SDs. *Protein precipitated during the reaction.

(< 0.8 M NaCl) and resulted in loss of activity (Fig. 2A). Though, almost one-third of the lost activity was regained when dialyzed against the high-salt buffer (2 M NaCl, 50 mM HEPES/NaOH, pH 7.5).

Importantly, ADH/A1a retained ~ 70–80% of its initial activity after 32 h in 10% (v/v) acetonitrile, methanol, and dimethylsulfoxide (Fig. 2B). At high solvent concentrations (30%, v/v), ADH/A1a was able to withstand dimethylsulfoxide and methanol and retained almost half of its initial activity after 32 h.

Structural analysis of ADH/A1a

The homology model of ADH/A1a was built based on the crystal structure of the thermophilic htADH of *Bacillus stearothermophilus* (PDB: 1rjw, 37% identity) [40]. The homology model of ADH/A1a appeared to

be reliable, with a QMEAN *Z*-score of −1.6 and a GMQE of 0.71. High structural conservation was found in the protein core domains, whereas a few less conserved areas were detected in the surface loops, indicating more structural variance (Fig. S10). The quaternary structure of ADH/A1a revealed a homotetramer bearing two conserved zinc-binding sites per unit (Fig. 3A). The catalytic binding site contains the coordinating residues cys42, his64, and cys151 and resembles the typical binding motif of the zinc-binding MDR superfamily (Fig. 3B). The structural zinc-binding site consists of four cysteine residues (cys95, cys98, cys101, and cys109) and is located in the interface region between two monomers. The identified N-terminal GroES-like domain encloses the catalytically active center and consists of four β-strands and a short α-helix that form a β-barrel-like structure [41]. Furthermore, the C-terminal nucleotide-binding domain is

composed of a classical Rossmann fold of six parallel beta-strands flanked by four alpha-helices [42].

The thermophilic htADH [40], the mesophilic FurX [43], and the psychrophilic MADH [44] are among the

Table 1. Substrate scope of ADH/A1a.

| Substrate	Relative activity (%)
(A) Substrates (0.2% v/v) in the oxidation reaction (100% = 0.49 µmol·min^{-1}·mg^{-1})	
Ethanol	15.2 ± 2.6
1-Propanol	11.9 ± 4.1
3-Buten-2-ol	9.2 ± 2.0
Benzyl alcohol	26.9 ± 2.0
Isoamyl alcohol	11.6 ± 2.1
Prenol	29.2 ± 2.0
1-Butanol	32.3 ± 5.5
1-Pentanol	46.5 ± 6.1
1,5-Pentanediol	59.6 ± 3.0
1-Heptanol	59.1 ± 1.2
Hydrocinnamyl alcohol	53.7 ± 1.6
Cinnamyl alcohol	100.0 ± 5.0
Geraniol	36.1 ± 2.4
Citronellol	33.5 ± 3.4
Nerol	14.9 ± 2.4
(B) Substrates (0.4% v/v) in the reduction reaction (100% = 0.40 µmol·min^{-1}·mg^{-1})	
Cinnamaldehyde	76.6 ± 4.3
Raspberry ketone	63.4 ± 5.2
Cinnamyl-methyl-ketone	100.0 ± 2.2

closest homologous structures (Table S3). A comparison with the homology model of ADH/A1a revealed high structural similarities (RMSDs: 0.63–0.7 Å; Fig. S11), although the sequence identity was comparably low (34–37%). Among these homologs, ADH/A1a exhibits the most negative electrostatic surface potential (Fig. S12), resulting from a great abundance of solvent-exposed acidic residues, particularly of glutamic acid (Figs S13 and S14). Similarly, but to a lesser extent, the cold-adapted MADH showed an increased negative surface charge, whereby the ratio of glutamic to aspartic acid is balanced. Moreover, both ADH/A1a and htADH exhibited increased amounts of charged surface-exposed residues, whereas the ratio of hydrophobic residues is reduced (Fig. S13).

Phylogenetic analysis of ADH/A1a

Alcohol dehydrogenase/A1a belongs to the CAD family of the MDRs [15]. The CDD divided the CAD family into four subfamilies (CAD1, CAD2, CAD3, and CAD-like) according to their evolutionarily conserved protein domain patterns by aiming to provide a functional characterization of distinct subfamilies [14]. ADH/A1a was annotated as a member of the subfamily CAD2, which has not been investigated to date. However, the functionality-based annotation by KEGG proposed the classification of ADH/A1a as a

Table 2. Kinetic parameters of ADH/A1a.

| Substrate	V_{max} (mU·mg^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}·mM^{-1})
Cinnamyl alcohol	236.8 ± 9.2	0.47 ± 0.09	0.144 ± 0.009	0.31 ± 0.01
Geraniol	130.9 ± 4.2	0.27 ± 0.04	0.082 ± 0.003	0.29 ± 0.018
Ethanol	142.1 ± 8.8	94.43 ± 8.4	0.089 ± 0.006	(0.94 ± 0.71) × 10^{-3}
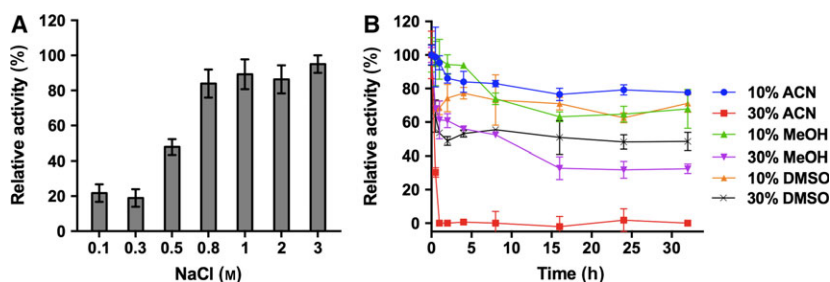
Fig. 2. Effect of salt concentration and solvent on the stability of ADH/A1a. (A) Activity of ADH/A1a after 24 h in various salt concentrations. ADH/A1a was dialyzed against varying salt concentrations (10 mM HEPES/NaOH, pH: 7.5, 0.1–3 M NaCl) at 4 °C for 24 h, and the remaining activity was determined. (B) Activity of ADH/A1a over time in different organic solvents. ADH/A1a (in 10 mM HEPES/NaOH, pH 7.5, 2 M NaCl) was incubated with different amounts of organic solvents at room temperature, and the residual activity was assayed over time. ADH/A1a incubated without solvent was treated as the control. The activities of ADH/A1a were assayed in the oxidation reaction under standard conditions. (100% = 0.49 µmol·min^{-1}·mg^{-1}). Error bars indicate SDs.

propanol-preferring ADH (EC 1.1.1.1, AdhP), including several suggested metabolic pathways. Notably, pathways of the classical CADs found in plants (EC 1.1.1.195) were not proposed. A sequence alignment of ADH/A1a with classical CADs and with propanol-preferring ADHs revealed that ADH/A1a shares features with both families (Fig. S15). Interestingly, a few features found in the sequence were unique to the CAD2 subfamily, as an extra loop at position 9–13 and a loop deletion at position 224–225. The structure

homologs, htADH, FurX, and MADH, were also classified as propanol-preferring ADHs (EC 1.1.1.1) by KEGG, but these ADHs belong to the CAD3 subfamily. CAD3 shares most of its features with ethanol active ADHs and is more closely related to CAD2 than to CAD1, based on our phylogenetic analysis. The phylogenetic tree reveals the positioning of the CAD2 subfamily between the CAD1 and the CAD3/ethanol active ADHs (Fig. 4). We found that the CAD2 subfamily comprises several uncharacterized homologous
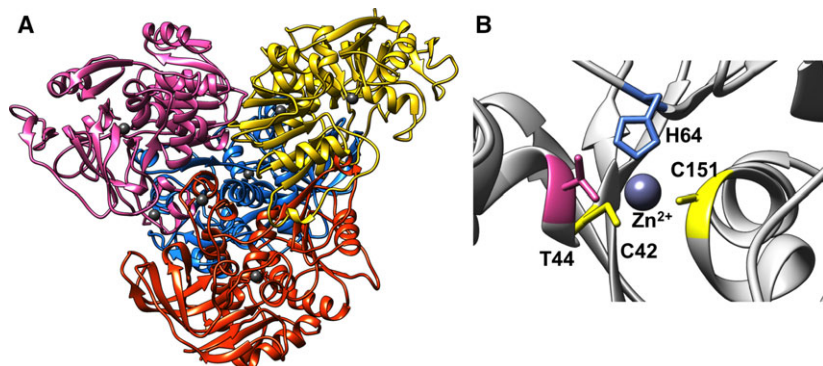
Fig. 3. Structure of the homology model of ADH/A1a. (A) Tetrameric structure of ADH/A1a. Monomers are indicated by different colors; Zn ions are dark gray. (B) Conserved catalytic Zn-binding motif cys42-his64-cys151 of ADH/A1a.
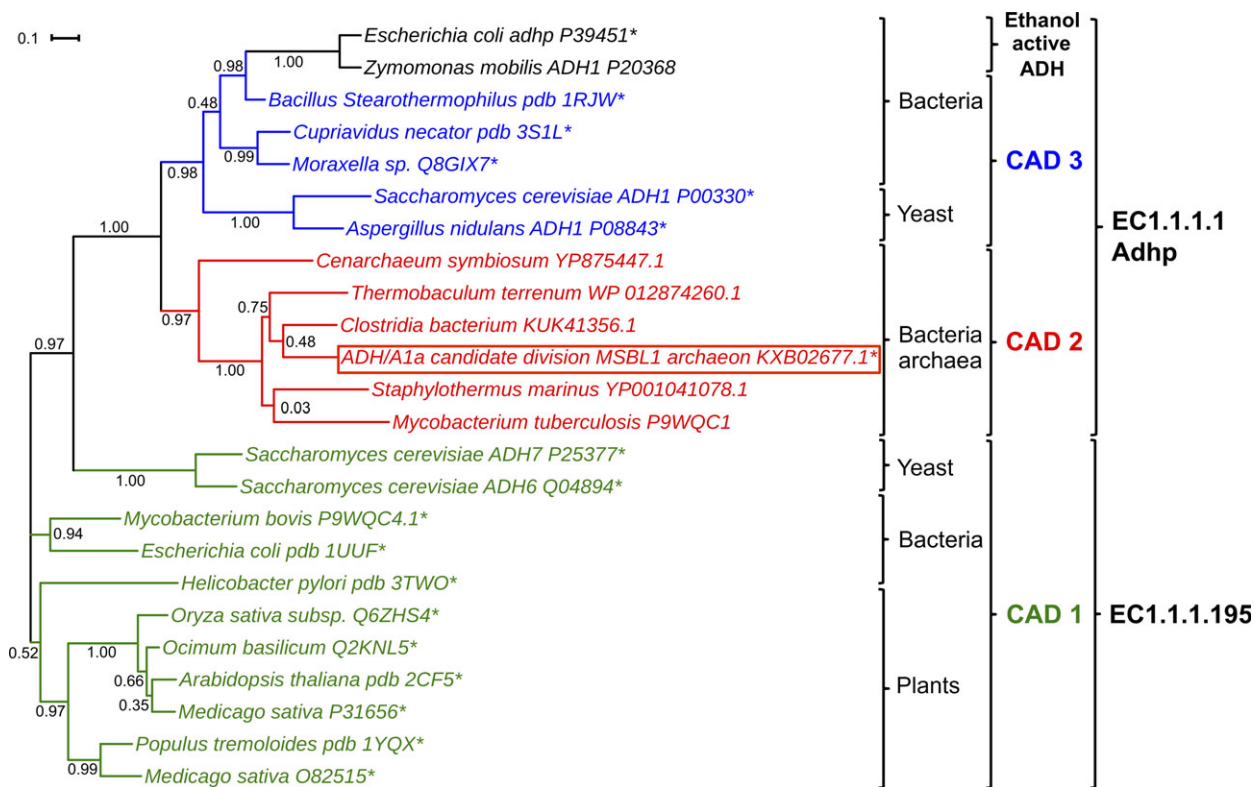
Fig. 4. Phylogenetic tree of ADH/A1a and selected ADHs of closest subfamilies. The branches are colored according to the subfamilies, which were annotated using CDD. The EC numbers indicate the functional annotation given by KEGG. The calculation of the phylogenetic tree was performed using the maximum-likelihood algorithm of FastTree2. *Experimentally characterized.

genes from genome annotations of bacteria and archaea, but not from plants or fungi. The investigation of these CAD2 sequences revealed that most of them were derived from extremophilic environments (Table S4). This evidence corroborates the uniqueness of the CAD2, whose representatives appear like propanol-preferring ADHs according to KEGG, but are able to convert substrates specific for both families, many of them being stable under extremophilic conditions.

Discussion

The Red Sea brine pools are an attractive source for the discovery of extremozymes. However, microorganisms inhabiting and adapted to these extreme environments are hardly cultivable. With the rise of culture-independent methods, this 'microbial dark matter' is now open for exploration [45]. We identified the ADH gene *adh/a1a* by utilizing annotated SAGs, derived from water samples of the Atlantis II Deep Red Sea brine pool [13]. Our investigation revealed that ADH/A1a is a polyextremophilic enzyme with special structural stabilities and functionalities, and further gained an understanding of its structural salt and temperature adaptations.

For characterization, we first attempted the expression of ADH/A1a in the mesophilic bacterium *E. coli*. However, ADH/A1a was not successfully expressed since the expression of halophilic enzymes is challenging and generally requires high intracellular salt concentrations [46,47]. Therefore, we used the halophilic *H. volcanii* expression system to express soluble and active ADH/A1a under high-salt conditions [17].

The extensive characterization of ADH/A1a revealed the halo-thermophilic character of the enzyme reflected by the activity optima at 4 M salt and 70 °C, resembling the natural conditions of the Atlantis II Deep brine pool. The preference for KCl over NaCl was expected, as KCl is the predominant cytoplasmic ion in cells [48]. The pH optima of the reactions performed with ADH/A1a were consistent with several reported ADHs [49]. The highly basic pH optimum of the oxidation reaction (pH 10) does not resemble the natural physiological conditions of a cell, but might be explained by the specific thermodynamics of the reaction. The enzyme equivalently converted the cofactors NAD(H) and NADP(H), despite the usual specificity of ADHs for one cofactor [50]. ADH/A1a was determined to be zinc dependent and required high zinc concentrations to reach optimal activity. This might be explained by the fact that the natural zinc concentration in the Atlantis II Deep sea brine pool is very high (165.7 μM) compared with normal sea water (0.2 μM)

[10]. Additionally, zinc ions could bind on the negatively charged surface of ADH/A1a or be involved in folding and structural stabilization of the enzyme [51,52]. Moreover, the his$_6$-tag on ADH/A1a and residual contaminants might bind zinc ions nonspecifically.

Alcohol dehydrogenase/A1a showed remarkable stability in the presence of high salt concentrations and elevated temperatures. However, low-salt conditions destabilized ADH/A1a, possibly because of the increased structural flexibility and partial unfolding promoted by the decline in external ionic forces [12]. Notably, the presence of salt is expected to play an important role in the structural protection of halophilic proteins [53]. Furthermore, ADH/A1a was able to withstand exposure to organic solvents. Mesophilic enzymes generally tend to lose their activity in the presence of organic solvents; however, halophilic enzymes are reported to be more stable [8,54]. The stability of ADH/A1a both in solvents and at high temperatures is valuable for biocatalytic reactions in various aspects, for example, increased solubility of organic compounds or improved reaction rates [55]. Slower reaction rates were observed compared to the reported homolog ADHs. However, this might be explained by the fact that the life in the extreme of the deep sea brine pools is probably slow and high enzyme activity of ADH is not crucial for survival. The adaptation to the extreme conditions of the habitat might have shifted toward increased stabilities but simultaneously also to reduced activities [21]. Furthermore, we cannot exclude that the non-native reaction environment and the halophilic host expression system might affect the overall native structure and function of the polyextremophilic ADH/A1a, resulting in reduced activity.

Despite the low sequence identity to homologs, we were able to build a reliable ADH/A1a homology model based on the significant structural similarity to the structures of homologous ADHs (RMSD below 1 Å). The enzyme core domains, including the active center and the nucleotide-binding site, are highly conserved, whereas the solvent-exposed residues and some flexible regions are altered, enabling the adaptation of the enzyme to the extreme conditions. SWISS-MODEL calculation and the native PAGE analysis suggest the homotetramer formation of ADH/A1a. Compared to nonhalophilic ADH homologs, the solvent-exposed surface of the ADH/A1a tetramer displayed increased amounts of glutamic acid residues, resulting in a high negative electrostatic surface potential. A negative surface charge is a known adaptation feature of halophilic enzymes and enables the binding of hydrated ions, which stabilize and increase the

solubility of the protein under high-salt/low-water conditions [8]. A similar adaptation mechanism that enhances enzyme solvation under cold conditions was reported in psychrophilic enzymes, as seen for MADH [56]. Furthermore, the surface of ADH/A1a exhibits an increased ratio of charged residues and a decreased ratio of hydrophobic residues, which probably diminishes hydrophobic patches on the surface and resembles another haloadaptive feature [48]. An increased ratio of charged amino acids on the surface was also reported for thermophilic enzymes, as we have seen for the homologous htADH [57]. The reduction of hydrophobic and the increase of charged residues on the surface add probably to both halo- and thermoadaptation.

Alcohol dehydrogenase/A1a effectively oxidized a broad spectrum of alcohols. In the reduction reaction, exclusively cinnamaldehyde, cinnamyl-methyl-ketone, and raspberry ketone were reduced. The substrate scope of ADH/A1a coincided with the substrates used by ADHs of the CAD family [58]. Based on its conserved domains, ADH/A1a was annotated as a member of the CAD2 subfamily, but its functionality was indicated to be of a propanol-preferring ADH, rather than of a classical CAD, perhaps due to the lack of experimentally verified CAD2 sequences. Phylogenetic analysis revealed the positioning of ADH/A1a between these ADH families, showing shared features with both, but also unique features specific for the CAD2 family. In plants, the function of CADs in lignin biosynthesis and plant defense mechanisms has been well characterized [59,60]; however, the role of CADs in microorganisms is not yet completely understood. The involvement in several metabolic pathways is considered including the Ehrlich pathway, NAD(P)$^+$/H homeostasis, lipid metabolism, amino acid metabolism, and lignin degradation [23,61,62]. ADH/A1a is the first ADH of the CAD2 subfamily that has been described; thus, its characterization may lead to further understanding of the metabolic functions. Aromatic substrate-preferring ADHs, such as CADs, are not well investigated to date, despite their great potential in the production of chiral aromatic precursors for pharmaceutical compounds [59], for example, health-related lignans as podophyllotoxin [63,64]. The substrate-specific reduction of aromatic ketones combined with the high stability makes ADH/A1a a possible candidate for the production of valuable chiral compounds.

Acknowledgements

Authors' contributions

JE and MR conceived and supervised the study; AA, RK, and JE designed the experiments; AA and RK performed experiments; AA, RK, AH, MV, and IA analyzed the data; SG provided plasmid and support; AA and RK wrote the manuscript; and AH, MV, IA, and SG made manuscript revisions.

Conflict of interest

The authors declare no conflict of interest.

References

1 Matsuda T, Yamanaka R and Nakamura K (2009) Recent progress in biocatalysis for asymmetric oxidation and reduction. *Tetrahedron: Asymmetry* **20**, 513–557.

2 Huisman GW, Liang J and Krebber A (2010) Practical chiral alcohol manufacture using ketoreductases. *Curr Opin Chem Biol* **14**, 122–129.

3 Zhang R, Xu Y and Xiao R (2015) Redesigning alcohol dehydrogenases/reductases for more efficient biosynthesis of enantiopure isomers. *Biotechnol Adv* **33**, 1671–1684.

4 Karan R and Khare SK (2010) Purification and characterization of a solvent-stable protease from Geomicrobium sp. EMB2. *Environ Technol* **31**, 1061–1072.

5 Eijsink VGH, Gåseidnes S, Borchert TV and van den Burg B (2005) Directed evolution of enzyme stability. *Biomol Eng* **22**, 21–30.

6 Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC and Robins K (2012) Engineering the third wave of biocatalysis. *Nature* **485**, 185–194.

7 Dumorne K, Cordova DC, Astorga-Eló M and Renganathan P (2017) Extremozymes: a potential source for industrial applications. *J Microbiol Biotechnol* **27**, 649–659.

8 Karan R, Capes MD and DasSarma S (2012) Function and biotechnology of extremophilic enzymes in low water activity. *Aquat Biosyst* **8**, 4.

9 Trincone A (2011) Marine biocatalysts: enzymatic features and applications. *Mar Drugs* **9**, 478.

10 Antunes A, Ngugi DK and Stingl U (2011) Microbiology of the Red Sea (and other) deep-sea anoxic brine lakes. *Environ Microbiol Rep* **3**, 416–433.

11 Sonbol SA, Ferreira AJS and Siam R (2016) Red Sea Atlantis II brine pool nitrilase with unique thermostability profile and heavy metal tolerance. *BMC Biotechnol* **16**, 14.

12 Takahashi M, Takahashi E, Joudeh LI, Marini M, Das G, Elshenawy MM, Akal A, Sakashita K, Alam I, Tehseen M *et al.* (2018) Dynamic structure mediates halophilic adaptation of a DNA polymerase from the deep-sea brines of the Red Sea. *FASEB J* **32**, 3346–3360.

13 Alam I, Antunes A, Kamau AA, Ba Alawi W, Kalkatawi M, Stingl U and Bajic VB (2013) INDIGO - Integrated data warehouse of microbial genomes with examples from the red sea extremophiles. *PLoS ONE* **8**, e82210.

14 Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR *et al.* (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* **45**, D200–D203.

15 Persson B, Hedlund J and Jörnvall H (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families: the MDR superfamily. *Cell Mol Life Sci* **65**, 3879–3894.

16 Stroud A, Liddell S and Allers T (2012) Genetic and biochemical identification of a novel single-stranded DNA-binding complex in *Haloferax volcanii*. *Front Microbiol* **3**, 224.

17 Strillinger E, Grotzinger SW, Allers T, Eppinger J and Weuster-Botz D (2016) Production of halophilic proteins using *Haloferax volcanii* H1895 in a stirred-tank bioreactor. *Appl Microbiol Biotechnol* **100**, 1183–1195.

18 Mwirichia R, Alam I, Rashid M, Vinu M, Ba-Alawi W, Anthony Kamau A, Kamanda Ngugi D, Goker M, Klenk HP, Bajic V *et al.* (2016) Metabolic traits of an uncultured archaeal lineage -MSBL1- from brine pools of the Red Sea. *Sci Rep* **6**, 19181.

19 Grotzinger SW, Alam I, Ba Alawi W, Bajic VB, Stingl U and Eppinger J (2014) Mining a database of single amplified genomes from Red Sea brine pool extremophiles-improving reliability of gene function prediction using a profile and pattern matching algorithm (PPMA). *Front Microbiol* **5**, 134.

20 Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC and Jahn D (2005) JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res* **33**, W526–W531.

21 Grotzinger SW, Karan R, Strillinger E, Bader S, Frank A, Al Rowaihi IS, Akal A, Wackerow W, Archer JA, Rueping M *et al.* (2018) Identification and experimental characterization of an extremophilic brine pool alcohol dehydrogenase from single amplified genomes. *ACS Chem Biol* **13**, 161–170.

22 Allers T (2008) *Haloferax volcanii* transformation - thorsten allers laboratory (Chapter 3.4). In *The Halohandbook: Protocols for Haloarchaeal Genetics* (Dyall-Smith M, ed.), Vol. 7, p. 1–144.

23 Wilkin J-M, Soetaert K, Stélandre M, Buyssens P, Castillo G, Demoulin V, Bottu G, Laneelle M-A, Daffe M and De Bruyn J (1999) Overexpression, purification and characterization of *Mycobacterium bovis* BCG alcohol dehydrogenase. *Eur J Biochem* **262**, 299–307.

24 Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI *et al.* (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res* **43**, D222–D226.

25 Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang H-Y, Dosztányi Z, El-Gebali S, Fraser M *et al.* (2017) InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res* **45**, D190–D199.

26 Drozdetskiy A, Cole C, Procter J and Barton GJ (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389–W394.

27 Arnold K, Bordoli L, Kopp J and Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201.

28 Roy A, Kucukural A and Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* **5**, 725–738.

29 Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L *et al.* (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**, W252–W258.

30 Benkert P, Kunzli M and Schwede T (2009) QMEAN server for protein model quality estimation. *Nucleic Acids Res* **37**, W510–W514.

31 Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G and Baker NA (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res* **35**, W522–W525.

32 Unni S, Huang Y, Hanson R, Tobias M, Krishnan S, Li WW, Nielsen JE and Baker NA (2011) Web servers and services for electrostatics calculations with APBS and PDB2PQR. *J Comput Chem* **32**, 1488–1491.

33 Guex N and Peitsch MC (1997) SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723.

34 Katoh K, Rozewicki J and Yamada KD (2017) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* **30**, 3059.

35 Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M *et al.* (2004) UniProt: the universal protein knowledgebase. *Nucleic Acids Res* **32**, D115–D119.

36 Price MN, Dehal PS and Arkin AP (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490.

37 Kanehisa M, Sato Y and Morishima K (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* **428**, 726–731.

38 Hering O, Brenneis M, Beer J, Suess B and Soppa J (2009) A novel mechanism for translation initiation operates in haloarchaea. *Mol Microbiol* **71**, 1451–1463.

39 Flamholz A, Noor E, Bar-Even A and Milo R (2012) eQuilibrator—the biochemical thermodynamics calculator. *Nucleic Acids Res* **40**, D770–D775.

40 Ceccarelli C, Liang Z-X, Strickler M, Prehna G, Goldstein BM, Klinman JP and Bahnson BJ (2004) Crystal structure and amide H/D exchange of binary complexes of alcohol dehydrogenase from *Bacillus stearothermophilus*: insight into thermostability and cofactor binding. *Biochemistry* **43**, 5266–5277.

41 Taneja B and Mande SC (1999) Conserved structural features and sequence patterns in the GroES fold family. *Protein Eng* **12**, 815–818.

42 Rossmann MG, Moras D and Olsen KW (1974) Chemical and biological evolution of nucleotide-binding protein. *Nature* **250**, 194–199.

43 Kang C, Hayes R, Sanchez EJ, Webb BN, Li Q, Hooper T, Nissen MS and Xun L (2012) Furfural reduction mechanism of a zinc-dependent alcohol dehydrogenase from *Cupriavidus necator* JMP134. *Mol Microbiol* **83**, 85–95.

44 Tsigos I, Velonia K, Smonou I and Bouriotis V (1998) Purification and characterization of an alcohol dehydrogenase from the Antarctic psychrophile *Moraxella sp.* TAE123. *Eur J Biochem* **254**, 356–362.

45 Rashid M and Stingl U (2015) Contemporary molecular tools in microbial ecology and their application to advancing biotechnology. *Biotechnol Adv* **33**, 1755–1773.

46 Timpson LM, Liliensiek AK, Alsafadi D, Cassidy J, Sharkey MA, Liddell S, Allers T and Paradisi F (2013) A comparison of two novel alcohol dehydrogenase enzymes (ADH1 and ADH2) from the extreme halophile *Haloferax volcanii*. *Appl Microbiol Biotechnol* **97**, 195–203.

47 Ortega G, Lain A, Tadeo X, Lopez-Mendez B, Castano D and Millet O (2011) Halophilic enzyme activation induced by salts. *Sci Rep* **1**, 6.

48 Danson MJ and Hough DW (1997) The structural basis of protein halophilicity. *Comp Biochem Physiol A Physiol* **117**, 307–312.

49 Loderer C, Wagner D, Morgenstern F, Spieß A and Ansorge-Schumacher MB (2018) Discovery of a novel thermostable Zn2+ -dependent alcohol dehydrogenase from *Chloroflexus aurantiacus* through conserved domains mining. *J Appl Microbiol* **124**, 480–490.

50 Weckbecker A, Groger H and Hummel W (2010) Regeneration of nicotinamide coenzymes: principles and applications for the synthesis of chiral compounds. *Adv Biochem Eng Biotechnol* **120**, 195–242.

51 Tannous E and Kanaya S (2014) Divalent metal ion-induced folding mechanism of RNase H1 from extreme halophilic archaeon *Halobacterium* sp. NRC-1. *PLoS ONE* **9**, e109016.

52 Tannous E, Yokoyama K, You D-J, Koga Y and Kanaya S (2012) A dual role of divalent metal ions in catalysis and folding of RNase H1 from extreme halophilic archaeon *Halobacterium* sp. NRC-1. *FEBS Open Bio* **2**, 345–352.

53 Sinha R and Khare SK (2014) Protective role of salt in catalysis and maintaining structure of halophilic proteins against denaturation. *Front Microbiol* **5**, 165.

54 Karan R, Singh SP, Kapoor S and Khare SK (2011) A novel organic solvent tolerant protease from a newly isolated *Geomicrobium* sp. EMB2 (MTCC 10310): production optimization by response surface methodology. *N Biotechnol* **28**, 136–145.

55 Adlercreutz P (2008) Fundamentals of biocatalysis in neat organic solvents. In *Organic Synthesis with Enzymes in Non-aqueous Media* (Carrera G and Riva S, eds), pp. 3–24. WILEY-VCH Verlag GmbH, Weinheim.

56 Siddiqui KS and Cavicchioli R (2006) Cold-adapted enzymes. *Annu Rev Biochem* **75**, 403–433.

57 Dominy BN, Minoux H and Brooks CL (2004) An electrostatic basis for the stability of thermophilic proteins. *Proteins* **57**, 128–141.

58 Mee B, Kelleher D, Frias J, Malone R, Tipton KF, Henehan GTM and Windle HJ (2005) Characterization of cinnamyl alcohol dehydrogenase of *Helicobacter pylori*. *FEBS J* **272**, 1255–1264.

59 Youn B, Camacho R, Moinuddin SG, Lee C, Davin LB, Lewis NG and Kang C (2006) Crystal structures and catalytic mechanism of the *Arabidopsis* cinnamyl alcohol dehydrogenases AtCAD5 and AtCAD4. *Org Biomol Chem* **4**, 1687–1697.

60 Bagniewska-Zadworna A, Barakat A, Lakomy P, Smolinski DJ and Zadworny M (2014) Lignin and lignans in plant defence: insight from expression profiling of cinnamyl alcohol dehydrogenase genes during development and following fungal infection in *Populus*. *Plant Sci* **229**, 111–121.

61 Larroy C, Rosario Fernández M, González E, Parés X and Biosca JA (2003) Properties and functional significance of *Saccharomyces cerevisiae* ADHVI. *Chem-Biol Interact* **143–144**, 229–238.

62 Pick A, Rühmann B, Schmid J and Sieber V (2013) Novel CAD-like enzymes from *Escherichia coli* K-12 as additional tools in chemical production. *Appl Microbiol Biotechnol* **97**, 5815–5824.

63 Beutner KR and von Krogh G (1990) Current status of podophyllotoxin for the treatment of genital warts. *Semin Dermatol* **9**, 148–151.

64 Canel C, Moraes RM, Dayan FE and Ferreira D (2000) Podophyllotoxin *Phytochemistry* **54**, 115–120.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Alignment of ADH/A1a and CAD2 consensus sequence.

Fig. S2. Size exclusion chromatography (SEC) purification of ADH/A1a.

Fig. S3. Coomassie-stained SDS/PAGE of purified ADH/A1a.

Fig. S4. Coomassie-stained native PAGE of purified ADH/A1a.

Fig. S5. Tryptic digest and LC-MS/MS analysis of ADH/A1a.

Fig. S6. Measured melting curve and fitting curve.

Fig. S7. Calculated free Gibbs energy $\Delta_r G'$ at different pH.

Fig. S8. Effect of various parameters on the ADH/A1a activity.

Fig. S9. Michaelis–Menten kinetic of various substrates.

Fig. S10. Quality assessment of the homology model.

Fig. S11. Alignment of tetrameric structures of ADH/A1a and homologs.

Fig. S12. Electrostatic surface charge of ADH/A1a tetramer and structure homologs.

Fig. S13. Composition of surface-exposed amino acids of ADH/A1a and homologs.

Fig. S14. Ratio of amino acids on the solvent-exposed surface.

Fig. S15. Sequence alignment of selected ADHs of the CAD1, CAD2, CAD3 subfamily and propanol-preferring ADHs.

Table S1. Identified contaminant proteins of the host strain.

Table S2. ICP-OES measurement of zinc concentration of an ADH/A1a solution.

Table S3. Information of closest homologous ADHs with known crystal structures.

Table S4. Information of closest ADHs based on amino acid sequence.