Article:
Miles, JA orcid.org/0000-0001-8839-9201, Machattou, P, Nevin-Jones, D et al. (4 more authors) (2019) Identification of a cyanobacterial aldehyde dehydrogenase that produces retinoic acid in vitro. Biochemical and Biophysical Research Communications, 510 (1). pp. 27-34. ISSN 0006-291X

https://doi.org/10.1016/j.bbrc.2018.12.171

Identification of a cyanobacterial aldehyde dehydrogenase that produces retinoic acid in vitro

Jennifer A. Miles [a,b], Petrina Machattou [a], David Nevin-Jones [a], Michael E. Webb [a,b], Andrew Millard [c], David J. Scanlan [d] and Paul C. Taylor [a,b]*

[a] School of Chemistry, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK

[b] Astbury Centre for Structural Molecular Biology, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK

[c] Department of Infection, Immunity and Inflammation, University of Leicester, Leicester, LE1 7RH, UK

[d] School of Life Sciences, University of Warwick, Coventry, CV4 7AL, UK

* To whom correspondence should be addressed

Email: p.c.taylor@leeds.ac.uk

Phone: +44 (0)113 343 6529

Paul Taylor is a Director of Tangent Reprofiling Limited. Other authors have no competing financial interests.

Abstract

Retinoic acid signalling is generally considered to be of animal origin. Recently, retinoic acid has been identified in cyanobacteria, yet no mechanism for its production has been identified. Here, we characterize for the first time a cyanobacterial aldehyde dehydrogenase that produces retinoic acid in vitro. Our computational studies suggest that the cyanobacterial aldehyde dehydrogenase resembles an ancestor of both eukaryotic aldehyde dehydrogenase 1 and aldehyde dehydrogenase 2. The Chlorogloeopsis fritschii aldehyde dehydrogenase described here may find applications in synthetic production of retinoic acid as well as contributing to our understanding of retinoid synthesis in cyanobacteria.

Keywords: Aldehyde dehydrogenase, cyanobacteria, lateral gene transfer**,** retinoic acid,

Abbreviations

ADH, alcohol dehydrogenase: ALDH, aldehyde dehydrogenase; cfALDH, Chlorogloeopsis fritschii Aldehyde dehydrogenase; CMO, Carotenoid monooxygenase; CYP, Cytochrome; LC-MS, Liquid chromatography–mass spectrometry; MW, Molecular weight; NAD, Nicotinamide adenine dinucleotide; RAR, Retinoic acid Receptor; RXR, Retinoid X Receptor; RALDH, Retinal dehydrogenase; RDH, Retinol dehydrogenase; SEC, substrate entry channel; SUMO, Small Ubiquitin-like Modifier.

Introduction

Retinoic acid is a key signalling molecule in healthy development and in differentiation of stem cells, albeit uncontrolled levels of retinoic acid can lead to mutagenesis [1]. In eukaryotes retinoic acid is made by oxidation of all-trans-retinal by retinal/aldehyde dehydrogenases (ALDH1A1, ALDH1A2, ALDH1A3, ALDH8) [2]. This step is preceded by the production of retinal from either retinol or beta-carotene (Figure 1a). As yet, this canonical signalling pathway has only been characterised in animals.

Retinoic acid has more recently been identified in cyanobacteria both grown in culture and in blooms isolated in China, where it is thought to be responsible for the mutagenesis of frogs [3–9] Thus far, no mechanism for cyanobacterial retinoic acid production has been defined, beyond it being a potential degradation product of retinal [10].

Orthologues of carotenoid dioxygenases, used by humans to produce retinal from beta-carotene, have previously been characterised from the cyanobacteria Synechocystis PCC 6803 and Nostoc PCC 7107 [11–14]. These orthologues are able to convert beta-carotene and apo-8-carotenal into all-trans-retinal. An orthologue of a cytochrome P450 enzyme known as CYP26 in animals and CYP120 in cyanobacteria has also been characterised and found to oxidise retinoic acid [15–17].

All-trans-retinal and the precursor beta-carotene have been identified in 24 species of cyanobacteria [18]. Production of all-trans-retinal, beta-carotene or retinoic acid in cyanobacteria occurs independently of whether the species is grown in culture or isolated from a bloom [4,18]. Therefore, given the conservation of genes involved in the synthetic pathway, one hypothesis is that the retinoic acid produced by cyanobacteria could be produced by the same biosynthetic pathway as in eukaryotes, potentially starting from beta-carotene (Figure 1a). **INSERT FIG1 HERE**

Attempts have previously been made to identify a cyanobacterial aldehyde dehydrogenase that is able to convert all-trans-retinal to retinoic acid, bridging the gap between the previous

characterised carotenoid dioxygenase and CYP120 [19]. This Synechocystis PCC 6803 slr0091 protein was able to convert long chain aldehydes into their respective acids but had no activity against retinal.

We have previously postulated that retinoic acid is not of animal origin as previously thought, but may be a product of bacterial origin and present in humans through a lateral gene transfer event of aldehyde dehydrogenase and CYP120 from cyanobacteria to animals [20]. Previously, lateral gene transfers between cyanobacteria and animals, in particular of cyclooxygenases [21], calpains [22] and WD40 domains have been suggested [23], indicating that these events are likely to have shaped processes across the animal kingdom.

Our phylogenetic analysis identified orthologues of human ALDH1A1 in cyanobacteria that have a very high sequence conservation (>60%) and retain the catalytic GQCC motif. Herein we characterise in vitro a cyanobacterial ALDH and demonstrate that it can convert retinal to retinoic acid.

Results

A good model in for the study of retinoic acid production in cyanobacteria would have orthologues of all the proteins in the putative pathway. Therefore, we searched the NCBI database of non-redundant protein sequences for a cyanobacterium that contains orthologues of human ALDH1A2 (BAA34785.1), beta-carotene monooxygenase (AAI26211.1) and cytochrome P450 CYP26A1 (AAB88881.1). One of 65 cyanobacteria identified that possess all three orthologues (Tables I and S1), was Chlorogloeopsis fritschii PCC 6912.

To confirm that the C. fritschii ALDH orthologue also appears to have been involved in lateral gene transfer from cyanobacteria to animals we undertook phylogenetic analysis. Sequences were chosen from [20] and [24], alongside the sequence of the previously studied Synechocystis sp. PCC 6803 ALDH (Slr0091) and ALDH from Bacillus subtilis which is the first bacterial aldehyde dehydrogenase to be identified that can convert retinal to retinoic acid [25]**.** Phylogenetic analysis indicated that the C. fritschii ALDH orthologue (henceforth cfALDH), which has 64% sequence identity with human ALDH1A1, groups with other previously identified cyanobacterial proteins (Figure 1b and Supplementary Figure 1). When the slr0091 protein from Synechocystis PCC 6803 is included in the phylogenetic analysis (Figure 1b), this protein does not group with any of the other cyanobacterial proteins but rather groups with human ALDH8. This is consistent with the lack of function seen with the purified slr0091 [19]. The Bacillus cereus ALDH in our phylogenetic analysis shows a closer relationship with yeast ALDH and mammalian ALDH8 then it does with the cyanobacterial ALDH [25]

Sequence alignment of cfALDH with the human ALDH1 indicates that the cyanobacterial protein contains the same GQCC motif as the human ALDH1 and ALDH2 proteins [24] (Figure 1d, highlighted in red and Supplementary Figure 2). This motif is present in ALDH1/2 orthologues and these residues reside at the bottom of the substrate entry channel (SEC). The second cysteine is involved in the nucleophilic attack of the carbonyl carbon in the aldehyde (Figure 1c).

To ascertain the activity of the protein, cfALDH was cloned from genomic DNA and over-expressed in E. coli. SDS-PAGE analysis showed the purified protein has a MW ~50 kDa (Figure 2a), with analysis by accurate mass spectrometry giving a mass of 53613.60 Da whilst the calculated mass is 53614.19 Da (Supplementary Figure 3). The protein purified as a homo-tetramer on size-exclusion chromatography (calculated mass of ~239 kDa, Figure 2b and calibration in Supplementary Figure 4). **INSERT FIG2 HERE**

To establish whether retinoic acid is produced by this putative retinal dehydrogenase, an LC-MS based method was used. The purified protein was incubated in the presence of NAD and all-trans-retinal to determine the conversion to retinoic acid. Samples were extracted into hexane and analysed by LC-MS. After incubation at 37℃ for 90 minutes ret inoic acid was clearly produced when all components were present, but not without inclusion of the cofactor NAD (Figure 4c and d, standards shown in supplementary figure 5).

A fluorescence based assay was then used to measure NADH production and therefore calculate the levels of retinoic acid production. The concentration of all-trans-retinal was varied between 0.1-10 µM whilst NAD^+ remained at a high concentration of 1mM, whilst cfALDH was kept low at 200 nM (Figure 3a). The concentration of cfALDH was varied and a relationship between activity and enzyme concentration is seen as expected (Figure 3b). This process was optimal at a basic pH of 8.5 - 9.5 (Figure 3c) and enhanced by the presence of Mg^{2+} in the assay buffer (Figure 3d), as seen with other ALDH1 [26–30]. (Supplementary figure 6). This is the first time a cyanobacterial ALDH has been seen to produce retinoic acid from all-trans-retinal. **INSERT FIG3 HERE**

Previous work identified signature residues found in either ALDH1 or ALDH2 that potentially enable the prediction of whether an ALDH is able to process large or small aldehydes [24]. When we compare the cfALDH sequence to these signatures, we find that the protein contains 8/34 signature residues of the sheep ALDH1A1 with which it has a 64% sequence identity (NP_001009778.1), 17/34 of human ALDH2 with which it has a 68% sequence identity (AAP36614.1) and 9/34 that match neither ALDH1A1/2 (Figure 4a and S7). These residues

were then mapped onto the predicted model of cfALDH, generated using Phyre2 v.2.0 [31], where they are present across the surface and core of cfALDH (Figure 4c). **INSERT FIG4 HERE**

We then measured the estimated size of the substrate entry channel (SEC) using CastP 3.0 [32]. This enabled us to assess whether the cyanobacterial ALDH was likely to be more characteristic of an ALDH1 or ALDH2 considering it appears to be equally related to these proteins from our phylogenetic analysis (Figure 1b). The size of the SEC in an ALDH has previously been suggested as a way of predicting the substrate [24,33]**.**

The SEC in cfALDH is predicted to have a volume between that seen in the sheep ALDH1A1 (Volume 670.2 $Å^3$) and the human ALDH2 (Volume 558.07 $Å^3$) with a calculated volume of 635.05 $Å^3$ when the sheep ALDH1A1 1BXS is used as the template (Figure 4b). The GQCC motif is at the bottom of the SEC in cfALDH as expected. This suggests that the cyanobacterial (and planctomyces) ALDH could resemble the evolutionary branching point ahead of the evolution of animal ALDH1/2 that were selective for large or small aldehydes, respectively.

Discussion

We have identified and characterised for the first time a cyanobacterial ALDH that can convert all-trans-retinal to retinoic acid. Blooms producing retinoic acid are proving toxic to local flora and fauna, so any further insight into the mechanism of this production is crucial. This is a potentially important step in elucidation of the biosynthesis of retinoic acid in cyanobacteria, though further experimentation is required to determine if this enzyme is in fact employed for this purpose in this cyanobacterium.

We then undertook computational studies that suggest the cyanobacterial ALDH (and planctomyces ALDH) represent the evolutionary ancestor of human and other eukaryotic ALDH1/2 given that it contains signature residues matching both proteins throughout the surface of the protein and the core, along cfALDH being predicted to have an intermediate sized substrate entry channel, which allows in all-trans-retinal. Our work also suggests that indeed the size of the SEC is useful in predicting whether an ALDH is able to function on large or small aldehydes [24].

Future work should focus on establishing whether this aldehyde dehydrogenase is responsible for retinoic acid production in vivo and defining the role of retinoic acid in cyanobacteria.

Finally, we note that improved methods for the laboratory synthesis of retinoic acid are in demand, since this molecule is needed for treatment of skin cancers and acne [34]**.** Bacterial enzymes have been proposed as important in this field, with work underway to produce a metabolically engineered E. coli that can convert glycerol all the way through to retinoic acid [35,36]**.** The identification of a cyanobacterial enzyme able to catalyse the ultimate step in this biosynthesis may contribute to these developments.

Materials and Methods

Phylogenetic analyses

Aldehyde dehydrogenase sequences were chosen from [20] and [24], along with the Bacillus subtillis ALDH from [25] and slr0091 from Synechocystis PCC 6803 [19]. Amino acid sequences were aligned using MUSCLE [37] in MEGA7 [38]. Phylogeny was inferred by Maximum Likelihood using the LG model [39] in MEGA7 with 1000 bootstrap replicates. Full details of sequences are in Supplementary Information Table 2.

Cloning, expression and purification cfALDH

A cyanobacterial ALDH from Chlorogloeopsis fritschii PCC 6912 was identified by BLASTp searching for bacterial orthologues of the human ALDH1A2 (Accession number O94788.3). Genomic DNA was extracted from Chlorogloeopsis fritschii PCC 6912 cells grown in BG11 medium [40] at 30°C with a light intensity of 20 µmol m^{-2} s^{-1} [41]**.** Primers were designed to amplify the ALDH orthologue (WP_016874813.1) from genomic DNA in preparation for cloning [42] into pET28a-SUMO expression vector with an N-terminal His-SUMO tag. The cfALDH construct in pET28a-SUMO was transformed into Rosetta cells (Roche) (sequence shown in Figure S7). Cultures were grown at 37℃ until an OD 600 nm of 0.6-0.8, with expression induced with 0.5 mM IPTG overnight at 18℃. Cells were pelleted before resuspension in 0.1 M TRIS pH 7.5, 0.5 M NaCl, 20 mM Imidazole, 250 µM TCEP and EDTA-free protease inhibitor tablets (Roche). Cells were lysed by sonication for 3 x 1 min at 70% power. Soluble extracts were retrieved at 24,000 g for 45 minutes. Soluble fraction was incubated with Ni-NTA equilibrated in 0.1 M Tris-HCl pH 7.5, 0.5 M NaCl, 250 µM TCEP, 20 mM Imidazole. Beads were washed with 150 mM KCl, 0.1 M Tris-HCl pH 7.5, 250 µM TCEP, 20 mM Imidazole (5 CV) and 1 M NaCl, 100 mM Tris-HCl pH 7.5, 250 µM TCEP, 20 mM Imidazole (5 CV) (adapted from [43]). Further washes were undertaken with 0.1 M Tris-HCl pH 7.5, 0.5 M NaCl, 80 mM Imidazole and protein eluted with 0.1 M Tris-HCl pH 7.5, 0.5 M NaCl, 250 µM TCEP, 300 mM Imidazole. cfALDH was dialysed with Ulp1 (Ubl-specific protease 1) overnight into 50 mM Tris-

HCl pH 7.5, 0.25 M NaCl, 250 µM TCEP. Uncut material was removed by rebinding to Ni-NTA and cfALDH was further purified by size exclusion chromatography on Superdex 200 26/60 or Superdex 200 16/60 into 50 mM Tris-HCl pH 7.5, 0.25 M NaCl, 250 µM TCEP. Protein was concentrated in a 10 kDa cut-off concentrator and stored in aliquots at -80℃.

cfALDH activity assay

NAD stock was prepared from solid fresh buffer before every assay (Roche). NADH was purchased from Sigma. All-trans-retinal (Sigma) was prepared in methanol and stored in amber vials at -20 ℃. The assay buffer used for enzyme kinetics was 50 mM HEPES pH 8.5, 150 mM KCl, 2 mM DTT, 20 mM $MgCl_2$, 1 mM EDTA. All buffers and components were pre-warmed to 37℃ before setting up the assay. Assays were undertaken in 96 well black plates (Greiner). Total assay volumes were 200 µl. Retinal was used at a concentration range of 0.1-10 µM, with NAD at 1 mM, methanol 5% (v/v) and cfALDH at 200 nM. When the concentration of cfALDH was altered the concentration of all-trans-retinal was kept at 100 µM, with NAD at 1mM. The assay was carried out using a Perkin Elmer EnVision™ Multilabel Plate Reader to assess the formation of NADH by monitoring fluorescence measurements with excitation of the samples at 340 nm and measuring the emission at 460 nm. Measurements were taken every 2.2 second and the initial rate for each substrate concentration was calculated from the slope of the reaction kinetics in the linear range of the reaction in the first 200 seconds. The fluorescence of NADH stocks of known concentration was used to create a calibration curve. Each concentration was repeated at least 3 times and each experiment was repeated twice.

Effect of pH

A range of pH's were tested to find the optimum for cfALDH activity. Between pH 6.0 - 7.0 MES buffer at 50 mM was used, with HEPES buffer 50 mM at pH 7.5 - 8.5 and CAPS buffer 50 mM at pH 9.0 - 10.0. Every buffer also contained 150 mM KCl, 2 mM DTT, 20 mM $MgCl_2$ and 1 mM EDTA. cfALDH was kept at 500 nM with 1 mM NAD and 100 µM retinal used.

Effect of magnesium chloride concentration

$MgCl_2$ was varied from 0 mM to 20 mM to test the requirement of this cofactor, with the buffer also containing 50 mM HEPES pH 8.5, 150 mM KCl, 2 mM DTT and 1 mM EDTA. The concentration of cfALDH was kept at 500 nM, with 1mM NAD and 100 µM all-trans-retinal in each experiment.

Analysis of retinoic acid production via LC-MS

The assay buffer used for Liquid Chromatography-Mass Spectrometry analysis was 50 mM HEPES pH 8.5, 150 mM KCl, 2 mM DTT, 20 mM $MgCl_2$, 1 mM EDTA. cfALDH was used at 20 µM with 1 mM NAD and 100 µM all-trans-retinal in methanol (5% v/v). Samples had a total volume of 500 µl and were incubated at 37℃ for 90 mins before the addition of 500 µl methanol. Products were then extracted twice in 500µl hexane with firstly 10 µl 0.1M NaOH then 40 µl 5 M HCl. Samples were recorded by HP-LCMS, which was generally carried out on an Agilent 1200 series LC system comprising a Bruker HCT Ultra ion trap mass spectrometer. The solvent system used was CH_3CN/H_2O + 0.1% (v/v) formic acid with a Phenomenex Luna C18 50 × 2 mm 5 micron column. Samples were compared to standards of all-trans-retinal and retinoic acid (Sigma), both at 0.5 mg/ml in methanol (Supplementary Figure 5).

Accurate mass spectrometry

CfALDH (1 ml, 40 µM) was prepared in water for accurate mass analysis (Bruker MicrOTOF), using mobile phase (acetonitrile/water, 0.1% formic acid), gradient (2 – 95% MeCN over 1 min) and flow rate (1.3 mL/min).

Modelling the structure of the Chlorogloeopsis fritschii ALDH

A model of the structure of cfALDH was generated using Phyre2 [31]. The template chosen was 1BXS, Sheep ALDH1A1 [33]. This model was analysed using CastP [32]. Structures for Human ALDH2 (PDB code 2VLE [44]**)** and Sheep ALDH1A1 (PDB code 1BXS [33]**)** were also analysed using CastP.

Acknowledgements

We thank Dr David Adams at the University of Leeds for providing us with Chlorogloeopsis fritschii PCC 6912. This phylogenetic work was undertaken on MARC1, part of the High Performance Computing and LIDA facilities at the University of Leeds, UK.

References

[1] M. Rhinn, P. Dolle, Retinoic acid signalling during development, Development. (2012). doi:10.1242/dev.065938.

[2] N.Y. Kedishvili, Enzymology of retinoic acid biosynthesis and degradation, J. Lipid Res. (2013). doi:10.1194/jlr.R037028.

[3] X. Wu, J. Hu, A. Jia, H. Peng, S. Wu, Z. Dong, Determination and occurrence of retinoic acids and their 4-oxo metabolites in Liaodong Bay, China, and its adjacent rivers, Environ. Toxicol. Chem. (2010). doi:10.1002/etc.322.

[4] X. Wu, J. Jiang, Y. Wan, J.P. Giesy, J. Hu, Cyanobacteria blooms produce teratogenic retinoic acids, Proc. Natl. Acad. Sci. (2012). doi:10.1073/pnas.1200062109.

[5] A. Jonas, V. Buranova, S. Scholz, E. Fetter, K. Novakova, J. Kohoutek, K. Hilscherova, Retinoid-like activity and teratogenic effects of cyanobacterial exudates, Aquat. Toxicol. (2014). doi:10.1016/j.aquatox.2014.06.022.

[6] J. Javůrek, E. Sychrová, M. Smutná, M. Bittner, J. Kohoutek, O. Adamovský, K. Nováková, S. Smetanová, K. Hilscherová, Retinoid compounds associated with water blooms dominated by Microcystis species, Harmful Algae. (2015). doi:10.1016/j.hal.2015.06.006.

[7] E. Sychrová, J. Priebojová, M. Smutná, K. Nováková, J. Kohoutek, K. Hilscherová, Characterization of total retinoid-like activity of compounds produced by three common phytoplankton species, Harmful Algae. (2016). doi:10.1016/j.hal.2016.11.002.

[8] J. Priebojová, K. Hilscherová, T. Procházková, E. Sychrová, M. Smutná, Intracellular and extracellular retinoid-like activity of widespread cyanobacterial species, Ecotoxicol. Environ. Saf. (2018). doi:10.1016/j.ecoenv.2017.12.048.

[9] M. Smutná, J. Priebojová, J. Večerková, K. Hilscherová, Retinoid-like compounds produced by phytoplankton affect embryonic development of Xenopus laevis, Ecotoxicol. Environ. Saf. (2017). doi:10.1016/j.ecoenv.2016.12.018.

[10] Q. Chen, J.B. van der Steen, J.C. Arents, A.F. Hartog, S. Ganapathy, W.J. de Grip, K.J. Hellingwerf, Deletion of sll1541 in Synechocystis sp. strain PCC 6803 allows formation of a far-red-shifted holoproteorhodopsin in vivo, Appl. Environ. Microbiol. (2018). doi:10.1128/AEM.02435-17.

[11] S. Ruch, P. Beyer, H. Ernst, S. Al-Babili, Retinal biosynthesis in Eubacteria: In vitro characterization of a novel carotenoid oxygenase from Synechocystis sp. PCC 6803, Mol. Microbiol. (2005). doi:10.1111/j.1365-2958.2004.04460.x.

[12] D. Scherzinger, S. Ruch, D.P. Kloer, A. Wilde, S. Al-Babili, Retinal is formed from apo-carotenoids in Nostoc sp. PCC7120: in vitro characterization of an apo-carotenoid oxygenase, Biochem. J. (2006). doi:10.1042/BJ20060592.

[13] H. Cui, Y. Wang, S. Qin, Genomewide analysis of carotenoid cleavage dioxygenases in unicellular and filamentous cyanobacteria, Comp. Funct. Genomics. (2012). doi:10.1155/2012/164690.

[14] J. Heo, S.H. Kim, P.C. Lee, New insight into the cleavage reaction of Nostoc sp. strain PCC 7120 carotenoid cleavage dioxygenase in natural and nonnatural carotenoids, Appl. Environ. Microbiol. (2013). doi:10.1128/AEM.00071-13.

[15] N. Ke, J. Baudry, T.M. Makris, M.A. Schuler, S.G. Sligar, A retinoic acid binding cytochrome P450: CYP120A1 from Synechocystis sp. PCC 6803, Arch. Biochem. Biophys. (2005). doi:10.1016/j.abb.2005.01.011.

[16] K. Kühnel, N. Ke, M.J. Cryle, S.G. Sligar, M.A. Schuler, I. Schlichting, Crystal structures of substrate-free and retinoic acid-bound cyanobacterial cytochrome P450 CYP120A1, Biochemistry. (2008). doi:10.1021/bi800328s.

[17] A. Alder, P. Bigler, D. Werck-Reichhart, S. Al-Babili, In vitro characterization of synechocystis cyp120a1 revealed the first nonanimal retinoic acid hydroxylase, FEBS J. (2009). doi:10.1111/j.1742-4658.2009.07224.x.

[18] X. Wu, J. Jiang, J. Hu, Determination and occurrence of retinoids in a eutrophic lake (Taihu Lake, China): Cyanobacteria blooms produce teratogenic retinal, Environ. Sci. Technol. (2013). doi:10.1021/es303582u.

[19] D. Trautmann, P. Beyer, S. Al-Babili, The ORF slr0091 of Synechocystis sp. PCC6803 encodes a high-light induced aldehyde dehydrogenase converting apocarotenals and alkanals, FEBS J. (2013). doi:10.1111/febs.12361.

[20] A. Millard, D.J. Scanlan, C. Gallagher, A. Marsh, P.C. Taylor, Unexpected evolutionary proximity of eukaryotic and cyanobacterial enzymes responsible for biosynthesis of retinoic acid and its oxidation, Mol. Biosyst. (2014). doi:10.1039/c3mb70447e.

[21] A.R. Brash, N.P. Niraula, W.E. Boeglin, Z. Mashhadi, An ancient relative of cyclooxygenase in cyanobacteria is a linoleate 10S-dioxygenase that works In tandem with a catalase-related protein with specific 10S-hydroperoxide lyase activity, J. Biol. Chem. (2014). doi:10.1074/jbc.M114.555904.

[22] N.D. Rawlings, Bacterial calpains and the evolution of the calpain (C2) family of peptidases, Biol. Direct. (2015). doi:10.1186/s13062-015-0095-0.

[23] C.P. Ponting, L. Aravind, J. Schultz, P. Bork, E. V. Koonin, Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer, J. Mol. Biol. (1999). doi:10.1006/jmbi.1999.2827.

[24] T.J.P. Sobreira, F. Marlétaz, M. Simões-Costa, D. Schechtman, A.C. Pereira, F. Brunet, S. Sweeney, A. Pani, J. Aronowicz, C.J. Lowe, B. Davidson, V. Laudet, M. Bronner, P.S.L. De Oliveira, M. Schubert, J. Xavier-Neto, Structural shifts of aldehyde

dehydrogenase enzyme... [Proc Natl Acad Sci U S A. 2010] - PubMed result, Proc. Natl. Acad. Sci. U. S. A. (2010). doi:10.1073/pnas.1011223108.

[25] S.H. Hong, H.P.T. Ngo, H.K. Nam, K.R. Kim, L.W. Kang, D.K. Oh, Alternative biotransformation of retinal to retinoic acid or retinol by an aldehyde dehydrogenase from Bacillus cereus, Appl. Environ. Microbiol. (2016). doi:10.1128/AEM.00848-16.

[26] W.H. Takahashi K, Magnesium stimulation of catalytic activity of horse liver aldehyde dehydrogenase. Changes in molecular weight and catalytic sites., J. Biol. Chem. (1980) 8206–8209.

[27] I. Gagnon, G. Duester, P. V. Bhat, Kinetic analysis of mouse retinal dehydrogenase type-2 (RALDH2) for retinal substrates, Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol. (2002). doi:10.1016/S0167-4838(02)00213-3.

[28] I. Gagnon, G. Duester, P. V. Bhat, Enzymatic characterization of recombinant mouse retinal dehydrogenase type 1, Biochem. Pharmacol. (2003). doi:10.1016/S0006-2952(03)00150-3.

[29] A. Sima, M. Parisotto, S. Mader, P. V. Bhat, Kinetic characterization of recombinant mouse retinal dehydrogenase types 3 and 4 for retinal substrates, Biochim. Biophys. Acta - Gen. Subj. (2009). doi:10.1016/j.bbagen.2009.09.004.

[30] Y. Shabtai, H. Jubran, T. Nassar, J. Hirschberg, A. Fainsod, Kinetic characterization and regulation of the human retinaldehyde dehydrogenase 2 enzyme during production of retinoic acid, Biochem J. (2016). doi:10.1042/BCJ20160101.

[31] K. LA, Et.al, Phyre2, Nat. Protoc. (2015). doi:Doi 10.1590/S0102-311x2008001100012.

[32] J. Dundas, Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, J. Liang, CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues, Nucleic Acids Res. (2006).

doi:10.1093/nar/gkl282.

[33] S.A. Moore, H.M. Baker, T.J. Blythe, K.E. Kitson, T.M. Kitson, E.N. Baker, Sheep liver cytosolic aldehyde dehydrogenase: The structure reveals the basis for the retinal specificity of class I aldehyde dehydrogenases, Structure. (1998). doi:10.1016/S0969-2126(98)00152-X.

[34] S.H. Hong, K.R. Kim, D.K. Oh, Biochemical properties of retinoid-converting enzymes and biotechnological production of retinoids, Appl. Microbiol. Biotechnol. (2015). doi:10.1007/s00253-015-6830-8.

[35] H.J. Jang, B.K. Ha, J. Zhou, J. Ahn, S.H. Yoon, S.W. Kim, Selective retinol production by modulating the composition of retinoids from metabolically engineered E. coli, Biotechnol. Bioeng. (2015). doi:10.1002/bit.25577.

[36] H.J. Jang, S.H. Yoon, H.K. Ryu, J.H. Kim, C.L. Wang, J.Y. Kim, D.K. Oh, S.W. Kim, Retinoid production using metabolically engineered Escherichia coli with a two-phase culture system, Microb. Cell Fact. (2011). doi:10.1186/1475-2859-10-59.

[37] R.C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res. (2004). doi:10.1093/nar/gkh340.

[38] S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets, Mol. Biol. Evol. (2016). doi:10.1093/molbev/msw054.

[39] S.Q. Le, O. Gascuel, An improved general amino acid replacement matrix, Mol. Biol. Evol. (2008). doi:10.1093/molbev/msn067.

[40] R. Rippka, Isolation and Purification of Cyanobacteria, Methods Enzymol. (1988). doi:10.1016/0076-6879(88)67004-2.

[41] K. Stucken, J. Ilhan, M. Roettger, T. Dagan, W.F. Martin, Transformation and conjugal transfer of foreign genes into the filamentous multicellular cyanobacteria (subsection V) Fischerella and Chlorogloeopsis, Curr. Microbiol. (2012). doi:10.1007/s00284-012-

0193-5.

[42] F. Van Den Ent, J. Löwe, RF cloning: A restriction-free method for inserting target genes into plasmids, J. Biochem. Biophys. Methods. (2006). doi:10.1016/j.jbbm.2005.12.008.

[43] R. Bchini, V. Vasiliou, G. Branlant, F. Talfournier, S. Rahuel-Clermont, Retinoic acid biosynthesis catalyzed by retinal dehydrogenases relies on a rate-limiting conformational transition associated with substrate recognition, in: Chem. Biol. Interact., 2013. doi:10.1016/j.cbi.2012.11.019.

[44] E.D. Lowe, G.Y. Gao, L.N. Johnson, M.K. Wing, Structure of daidzin, a naturally occurring anti-alcohol-addiction agent, in complex with human mitochondrial aldehyde dehydrogenase, J. Med. Chem. (2008). doi:10.1021/jm800488j.

[45] A. Moretti, J. Li, S. Donini, R.W. Sobol, M. Rizzi, S. Garavaglia, Crystal structure of human aldehyde dehydrogenase 1A3 complexed with NAD+and retinoic acid, Sci. Rep. (2016). doi:10.1038/srep35710.

Figure Legends

Figure 1. (a) Retinoic acid production from beta-carotene or retinol. CMO - Carotenoid mono-oxygenase, RDH – Retinol dehydrogenase. ADH – Alcohol dehydrogenase. RALDH – Retinal dehydrogenase. RAR – Retinoic acid receptor. RXR – Retinoid X receptor, CYP26 - cytochrome P450 family 26, CYP120 - cytochrome P450 family 10. (b) Abridged phylogenetic tree of aldehyde dehydrogenase orthologues. Bootstrap values over 60% are shown. The full tree is shown in Figure S1. (c) Mechanism of aldehyde dehydrogenase catalysis. (d) Alignment of Chlorogloeopsis fritschii PCC 6912 ALDH (cfALDH WP_016874813.1) with human ALDH1A2 (O94788.3). The sequence identity is 64%. The GQCC motif is highlighted in red. The full alignment is shown in Figure S2.

Figure 2. (a) SDS-PAGE analysis of purified cfALDH. (b) Size-exclusion chromatography of purified cfALDH on Hiload Superdex 200 16/60. c) LC-MS analysis indicating that retinoic acid is produced when cfALDH is incubated with retinal and NAD. cfALDH incubated with all-trans-retinal, indicating that no retinoic acid is produced without the cofactor NAD (MW all-trans-retinal 284.4). d) cfALDH incubated with 1mM NAD and all-trans-retinal, indicating that retinoic acid (MW 300.4) is produced.

Figure 3. (a) Change in fluorescence seen when the concentration of all-trans-retinal is altered in the presence of cfALDH and NAD. (b) Change in retinoic acid (RA) production when the concentration of cfALDH is altered in the presence of all-trans-retinal and NAD. c) The effect of pH on cfALDH activity with the initial rate shown in µmol/mg/min. The activity is higher between pH 8.5-9.5. d) The effect of Mg^{2+} ion concentration on cfALDH activity with the initial rate shown in nmol/mg/min, with 20 mM giving the highest activity of the concentrations tested.

Figure 4. a) Identification of signature residues in cfALDH, with Sheep ALDH1A1 signature residues shown in red, ALDH2 in blue and residues matching neither signature in yellow [24]**.** b) Measurement of cavity size by CastP [32] in cfALDH compared to human ALDH2 (2VLE) [44] and Sheep ALDH1A1 (1BXS) [33]**.** The cfALDH structure was modelled using Phyre2 [31]

with 1BXS as a template [33]. c) Signature residues mapped onto the modelled structure of cfALDH. ALDH1A1 signature residues shown in red, ALDH2 in blue. Positions of NAD (grey) and retinal (yellow) were also modelled in by overlaying with the structure of human ALDH1A3 bound to NAD and all-trans-retinal (PDB code 5FHZ [45]).