Journal Name

Dynamic Article Links ►

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

# Unexpected Evolutionary Proximity of Eukaryotic and Cyanobacterial Enzymes responsible for Biosynthesis of Retinoic Acid and its Oxidation

**Andrew Millard,[a] David J. Scanlan,*[a] Connor Gallagher,[b] Andrew Marsh[b] and Paul C. Taylor*[b]**

***Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX***

**DOI: 10.1039/b000000x**

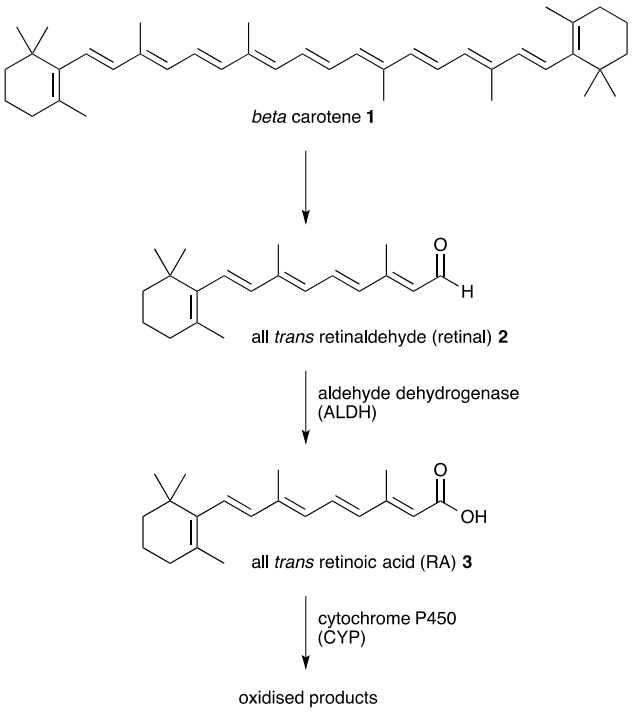
**Biosynthesis of retinoic acid from retinaldehyde (retinal) is catalysed by an aldehyde dehydrogenase (ALDH) and its oxidation by cytochrome P450 enzymes (CYPs). Herein we show by phylogenetic analysis that the ALDHs and CYPs in the retinoic acid pathway in animals are much closer in evolutionary terms to cyanobacterial orthologs than would be expected from the standard models of evolution.**

Retinoic acid (RA) is the main biologically active form of Vitamin A[1] and controls key development processes in the animal kingdom.[2] Indeed, RA signalling has been considered to be a distinctive feature of animals[3] with phylogenetic analyses pushing the evolutionary origins of retinoid biosynthesis ever earlier through the chordate[4] and bilaterian[5] lineages. RA is also medically important. Both RA and analogues are used as therapeutic drugs;[6] RA can be a harmful teratogen in other circumstances.[7]

Given its animal "heritage", it is surprising that RA, its biosynthetic precursor retinaldehyde and oxidised catabolites of RA have recently been discovered in cyanobacteria,[8] since animals and cyanobacteria are, in standard models, linked only at the very beginning of evolutionary history.[9] Research has focused little on the possibility that RA is a signalling molecule in cyanobacteria, with more emphasis on its seemingly aggressive teratogenic effects on other organisms during cyanobacterial blooms in, for example, eutrophic lakes in China.[10]

Our interests in sensory processes in cyanobacteria[11] and eukaryotic signalling molecules[12] led us to consider the possibility that there were hitherto unnoticed links between RA signalling in the animal kingdom and in cyanobacteria. Since lateral gene transfer between cyanobacteria and animals has been implied previously,[13] we reasoned it unwise to assume that RA signalling had evolved independently in these apparently utterly diverse lineages. Herein we extend to cyanobacteria phylogenetic analyses of the distinctive enzymes responsible for retinoid biosynthesis in animals and conclude that assumptions about separate evolution of RA pathways should be re-evaluated.

In vertebrates, retinoic acid **3** is biosynthesised from dietary retinal **2** (or dietary beta-carotene **1** via retinal **2**).[3] There are two enzymes that are highly characteristic of retinoid biosynthesis and catabolism and that have been the subjects of detailed phylogenetic analysis in eukaryotes, namely the aldehyde dehydrogenase (ALDH) that catalyses dehydrogenation of retinal **2** to form RA[5,14] and the cytochrome P450 that further transforms, or catabolises, RA.[15] Our study sought, very simply, to extend these phylogenetic analyses to prokaryotes.




In 2011 Sobreira *et al.* published their phylogenetic analysis of ALDH enzymes in eukaryotes, which included a range of organisms from across the evolutionary spectrum.[14] We submitted the protein sequence of human ALDH1A1 (EAW62544) to NCBI BLASTp[16] using standard settings but including only prokaryotes. Intriguingly, the closest alignments from all bacteria and archaea were with aldehyde dehydrogenases from cyanobacteria and from planctomycetes. We therefore added representative examples of these classes to the eukaryotic dataset used by Sobreira *et al*.[14] ALDH sequences were aligned in MUSCLE and viewed in MEGA5 (see Supplementary Information). Phylogeny was inferred by Maximum Likelihood using the Jones-Taylor-Thornton (JTT) model with 1000 bootstrap replicates. Our unrooted phylogenetic tree including cyanobacterial and planctomycete orthologs is shown in Figure 1.

According to Sobreira *et al.*, ALDH1a and ALDH2 enzymes fall into a common clade, in which the former specifically do and the latter specifically do not process retinaldehyde.[14] Our research

suggests that the bacterial orthologs should be considered phylogenetically as part of the same clade and that animal ALDH1a/2 enzymes are more closely related in evolutionary terms to these bacterial ALDHs than they are to their closest paralogs in the same animals, ALDH1L and ALDH8 dehydrogenases (Fig. 1).

The next biosynthetic step is the oxidation of RA by cytochrome P450 enzymes known as CYP26 in animals[6] and CYP120 in cyanobacteria.[17,18] A 2009 phylogenetic analysis of CYP26s by Albalat and Canestro gave an early clue that RA pathways may be 'ancient elements of animal genomes, already present in the last common ancestor of bilaterians'.[5] Once again, we first verified that we could reproduce Albalat and Canestro's tree using our methods. We then used NCBI BLASTp against the prokaryotic proteome to identify non-eukaryotic orthologs and, once again, found the best matches to be in cyanobacteria (though not planctomycetes).

Figure 2 presents our expanded phylogenetic tree for metazoan CYP26, plant CYP707 and cyanobacterial CYP120 orthologs. Animal CYP51 and CYP4v2 paralogs were added and the tree was rooted on human CYP4v2 as in the literature.[5] Again, the cyanobacterial CYP120s fall clearly within the clade including animal CYP26s and plant CYP707s

The two phylogenetic trees presented here are together strong evidence that the retinoic acid biosynthetic pathway outlined in Scheme 1 did *not* evolve separately in eukaryotes and prokaryotes. Rather the occurrence of two enzymes in the same pathway that both have structural and functional analogies across cyanobacteria and metazoans is more easily explained by a hitherto unsuspected lateral transfer of the corresponding genes.[19] Lateral gene transfer from cyanobacteria to animals has precedent; Aravind *et al.* have concluded that WD40 domains in metazoans most likely arose through such an event.[13,20]

Structural, spectroscopic and analytical studies show that CYP120A1 from *Synechocystis* and metazoan CYP26s are strikingly similar. They have both been shown to catalyse reactions of RA and to oxidise it at the 4- and 18- positions (oxidation at the 2-position has also been proposed in the cyanobacterium).[6,17,21,22] Furthermore, there are conserved tryptophan and phenylalanine residues believed to occupy important positions in the binding site. Indeed, Alder *et al.* have observed that 'despite the huge evolutionary distance' vertebrate and cyanobacterial retinoic acid hydroxylases are similar.[18] Our proposal of lateral gene transfer may resolve this conundrum.

To our knowledge there are no structural or mechanistic studies on the bacterial ALDH enzymes included in our phylogenetic analysis and the few annotations as "retinal dehydrogenase" appear to be from homology only. Therefore, we undertook a detailed study of the active site motif identified by Sobreira *et al.* as being characteristic of the catalytic site of these enzymes in vertebrates, GQCC, which is found at amino acids 299-302 of human ALDH1A1.[14]

The GQCC motif is found in many of the eukaryotic ALDH1 and ALDH2 enzymes in our study and in none of the ALDH8 and ALDH1L proteins. Strikingly the same motif is present in all of the bacteria we investigated, as indicated in Figure 1. This conserved motif at the catalytic site lends further support to our proposal that the similarities between animal and bacterial ALDHs are unlikely to have arisen through independent, convergent evolution.

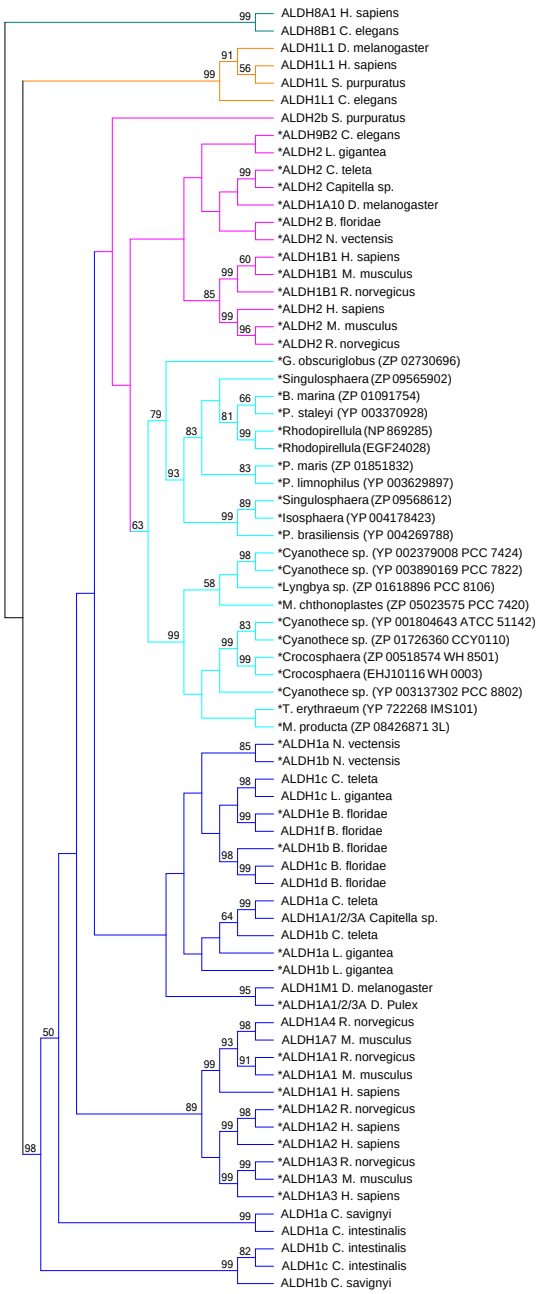



**Fig. 1** Phylogenetic tree of aldehyde dehydrogenase (ALDH) enzymes from representative metazoans, cyanobacteria and planctomycetes. Proteins analysed previously in Ref. 14 are shown in teal (ALDH8), orange (ALDH1L), blue (ALDH1) and pink (ALDH2). Cyanobacterial and planctomycete proteins added in this study are shown in cyan. Proteins with active site motif GQCC are indicated with an asterisk. In Fig. S1 (ESI) branches present in less than 50% of bootstrap replicates are collapsed.

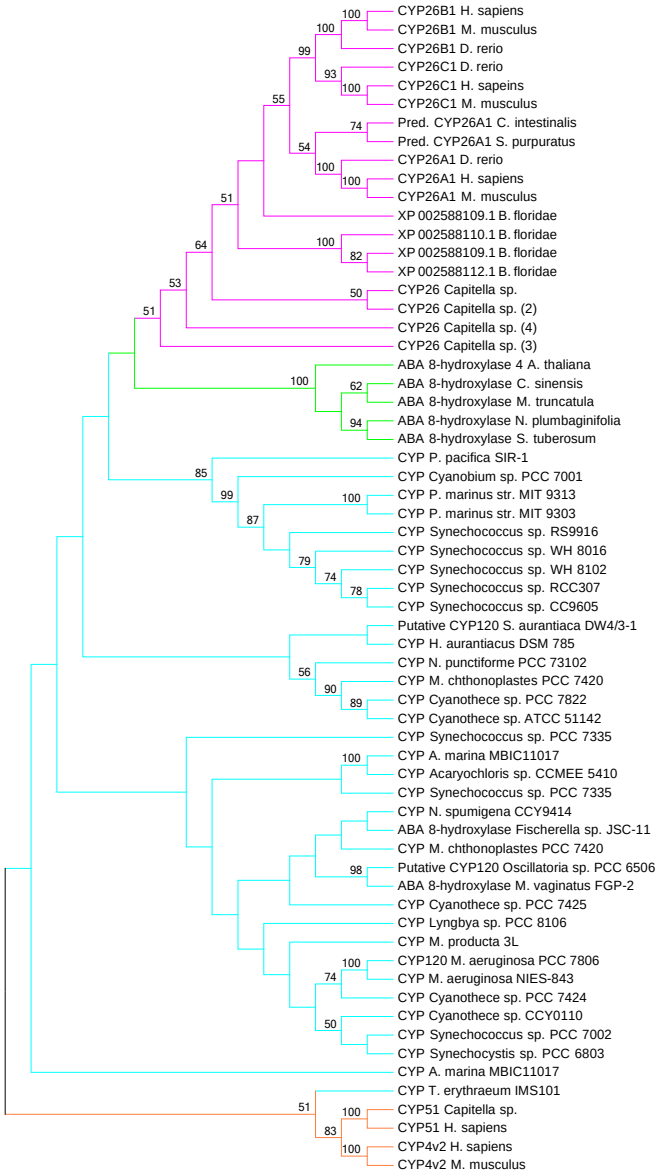



**Figure 2.** Phylogenetic tree of cytochrome P450 (CYP) enzymes from representative metazoans, plants and cyanobacteria. Proteins analysed previously in Ref. 5 are shown in pink (metazoan CYP26), orange (metazoan CYP51 & CYP4v2) and green (plant CYP707). Cyanobacterial proteins added in this study are shown in cyan. In Fig. S2 (ESI) branches present in less than 50% of bootstrap replicates are collapsed.

## Conclusions

Maximum likelihood phylogenetic analyses, *across all kingdoms*, of both ALDH1a/2 aldehyde dehydrogenases and cytochrome P450 enzymes known to oxidise retinoic acid show surprising evolutionary proximities between animal and cyanobacterial proteins (and also planctomycetes in the case of ALDHs). We believe the simplest explanation consistent with these data, and with literature observations on the structures and functions of the enzymes, to be a lateral gene transfer event between the bacteria and animals.

Given the importance of retinoic acid biosynthesis to developmental biology and in medicine, our observations suggest that detailed research of the pathway in cyanobacteria could have an impact on human healthcare,[2] as well as understanding of important ecosystems.[10] Further studies of the evolutionary links between such apparently distant species may also be revealing.

## Notes and references

[a] *School of Life Sciences, University of Warwick, Coventry, CV4 7AL, UK.*
[b] *Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK. Fax: 44 24 76524112; Tel: 44 24 76524375; E-mail: p.c.taylor@warwick.ac.uk*

† Electronic Supplementary Information (ESI) available: Methods. Accession numbers for proteins added in this study. Full names of organisms added in this study. Full page versions of Figs. 1, 2, S1 & S2. See DOI: 10.1039/b000000x/

1 J. Gutierrez-Mazariegos, M. Theodosiou, F. Campo-Paysaa, M. Schubert, *Semin. Cell Dev. Biol.*, 2011, **22**, 603.
2 M. Theodosiou, V. Laudet and M. Schubert, *Cell Mol. Life Sci.*, 2010, **67**, 1423.
3 R. Blomhoff and H. K. Blomhoff, *J. Neurobiol.*, 2006, **66**, 606.
4 M. S. Simoes-Costa, A. P. Azambuja and J. Xavier-Neto, *J. Exp. Zool. (Mol. Dev. Evol.)*, 2008, **310B**, 54.
5 R. Albalat and C. Canestro, *Chem. Biol. Interact.*, 2009, **178**, 188.
6 V. C. O. Njar, L. Gediya, P. Purushottamachar, P. Chopra, T. S. Vasaitis, A. Khandelwal, J. Mehta, C. Huynh, A. Belosay and J. Patel, *Bioorg. Med. Chem.*, 2006, **14**, 4323.
7 S.V. Bryant and D. M. Gardiner, *Dev. Biol.*, 1992, **152**, 1.
8 X-Q. Wu, J-Q. Jiang, Y. Wan, J. P. Giesy and J-Y. Hu, *Proc. Natl. Acad. Sci. U.S.A.*, 2012, **109**, 9477.
9 F. D. Ciccarelli, T. Doerks, C. von Mering, C. J. Creevey, B. Snel and P. Bork, *Science*, 2006, **311**, 1283.
10 X-Q. Wu, J-Q. Jiang and J-Y. Hu, *Environ. Sci. Technol.*, 2013, **47**, 807.
11 D. J. Scanlan, M. Ostrowski, S. Mazard, A. Dufresne, L. Garczarek, W. R. Hess, A. F. Post, M. Hagemann, I. Paulsen and F. Partensky, *Microbiol. Mol. Biol. Rev.*, 2009, **73**, 249.
12 S. J. Dilly, M. J. Bell, A. J. Clark, A. Marsh, R. M. Napier, M. J. Sergeant, A. J. Thompson and P. C. Taylor, *Chem. Commun.*, 2007, 2808.
13 L. Aravind, L. M. Iyer, V. Anantharaman. Natural History of Sensor Domains in Bacterial Signalling Systems in 'Sensory Mechanisms in Bacteria: Molecular Aspects of Signal Recognition', eds. Spiro S and Dixon R, Caister, UK, 2010.
14 T. J. P. Sobreira, F. Marlétaz, M. Simoes-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 and J. Xavier-Neto, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, **108**, 226.
15 A. R. Topletz, J. E. Thatcher, A. Zelter, J. D. Lutz, S. Taya, W. L. Nelson, N. Isoherranen, *Biochem. Pharm.*, 2012, **83**, 149.
16 S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller W and D. J. Lipman, *Nucleic Acids Res.*, 1997, **25**, 3389.
17 K. Kühnel, N. Ke, M. J. Cryle, S. G. Sligar, M. A. Schuler and I. Schlichting, *Biochemistry*, 2008, **47**, 6552.
18 A. Alder, P. Bigler, D. le Werck-Reichhart and S. Al-Babili, *FEBS Journal*, 2009, **276**, 5416.
19 M. A. Ragan and R. G. Beiko, *Phil. Trans. R. Soc. B.* 2009, **364**, 2241.
20 C. P. Ponting, L. Aravind, J. Schultz, P. Bork and E. V. Koonin, *J. Mol. Biol.*, 1999, **289**, 729.
21 M. S. Gomaa, S. W. Yee, C. E. Milbourne, M. C. Barbera, C. Simons and A. Brancale, *J. Enzyme Inhib. Med. Chem.*, 2006, **21**, 361.
22 K. Kaya, F. Shiraishi, H. Uchida and T. Sano, *Biochim. Biophys. Acta*, 2011, **1810**, 414.
23 A. L. Lamb and M. E. Newcomer, *Biochemistry*, 1999, **38**, 6003.