J Mol Evol (2002) 54:416–423
DOI: 10.1007/s00239-001-0031-z

JOURNAL OF MOLECULAR EVOLUTION

The Correct Phylogenetic Relationship of KdsA (3-Deoxy-D-*manno*-octulosonate 8-Phosphate Synthase) with One of Two Independently Evolved Classes of AroA (3-Deoxy-D-*arabino*-heptulosonate 7-Phosphate Synthase)

Roy A. Jensen,[1,2,3] Gary Xie,[2,3] David H. Calhoun,[1] Carol A. Bonner[2]

[1] Department of Chemistry, City College of New York, New York, NY 10031, USA
[2] Department of Microbiology and Cell Science, Gainesville, FL 32611, USA
[3] BioScience Division, Los Alamos National Laboratory, Los Alamos, NM 87544, USA

Received: 25 June 2001 / Accepted: 27 September 2001

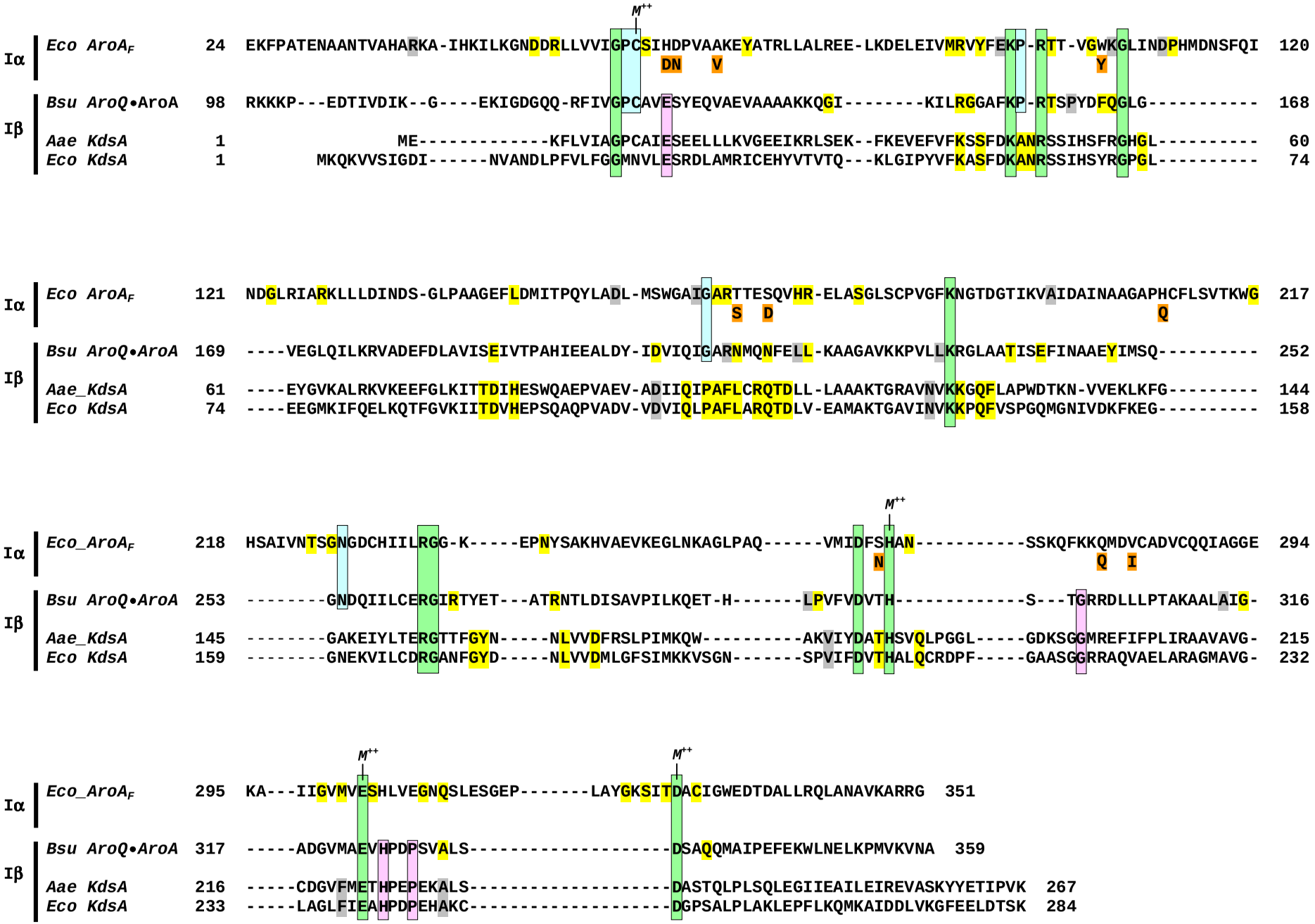
Recently Birck and Woodard (2001) published a phylogenetic tree of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthase proteins that not only incorrectly includes an unrelated class of DAHP synthase, but also incorrectly excludes 3-deoxy-D-*manno*-octulosonate 8-phosphate (KDOP) synthases.

DAHP is the initial product that is specifically committed to the biosynthesis of aromatic amino acids and a variety of other aromatic compounds via the action of DAHP synthase. KDOP is best known as a key precursor of lipopolysaccharide in gram-negative bacteria, but its wider distribution in capsular polysaccharides, lipoglycan of *Chlorella*, and cell walls of higher plants implies other functional roles at the cell surface as well [see Brabetz et al. (2000) and references therein]. In 1996 Walker et al. reported the existence of a class of DAHP synthase (their Class II) that lacked observable homology with the then-known DAHP synthases (their Class I). Class II DAHP synthases were described as 54,000-M_r enzymes that were homologues of higher-plant DAHP synthases, whereas Class I DAHP synthases (exemplified by the three well-known *Escherichia coli* paralogues) were 39,000-M_r enzymes. Class I DAHP synthases belong to the $AroA_I$ family of 3-deoxy-ald-2 ulosonate-phosphate synthases as defined by Subramaniam et al. (1998). Gosset et al. (2001) have recently produced an expanded analysis of the Class II DAHP synthases (denoted as the $AroA_{II}$ family), showing that the higher-plant sequences form a compact subcluster that is nested within the much more divergent set of microbial enzymes.

Although Birck and Woodard (2001) placed $AroA_{I\alpha}$ and $AroA_{II}$ sequences together, homology between the $AroA_I$ and $AroA_{II}$ families is not detectable at the level of primary sequence and cannot be asserted at the present time. It is possible that future tertiary structural information about $AroA_{II}$ might reveal a similar order and spacing of active-site residues that could be interpreted as evidence for insertion of loops in $AroA_{II}$ during divergent evolution, but structural data on $AroA_{II}$ are currently lacking. It is perhaps suggestive that the motifs GxC, KPR, and DxxH can be located as invariant residues in the $AroA_{II}$ family in the same order as present in the $AroA_I$ family (see Fig. 1). However, even if we knew that $AroA_I$ and $AroA_{II}$ are remote relatives, manual matching of amino acid residues based on criteria of colinear arrangement of amphiphilic patterns and critical spacing of motif patterns would be required to get a multiple alignment that would be satisfactory input for obtaining a tentative tree.

In 1998 Subramaniam et al. published an analysis indicating that the $AroA_I$ family of DAHP synthases ($AroA_I$) separates into two related subfamilies: $AroA_{I\alpha}$ (exemplified by the *E. coli* paralogues) and $AroA_{I\beta}$ (ex-

Correspondence to: Roy A. Jensen; *email:* rjensen@ufl.edu


Iα
Eco AroA_F 24 EKFPATENAANTVAHARKA-IHKILKGNDDRLLVVIGPCSIHDPVAAKEYATRLLALREE-LKDELEIVMRVYFEKP-RTT-VGWKGLINDPHMDNSFQI 120
Iβ
Bsu AroQ•AroA 98 RKKKP---EDTIVDIK--G----EKIGDGQQ-RFIVGPCAVESYEQVAEVAAAAKKQGI--------KILRGGAFKP-RTSPYDFQGLG---------- 168
Aae KdsA 1 ME------------KFLVIAGPCAIESEELLLKVGEEIKRLSEK--FKEVEFVFKSSFDKANRSSIHSFRGHGL---------- 60
Eco KdsA 1 MKQKVVSIGDI------NVANDLPFVLFGGMNVLESRDLAMRICEHYVTVTQ---KLGIPYVFKASFDKANRSSIHSYRGPGL---------- 74
Iα
Eco AroA_F 121 NDGLRIARKLLLDINDS-GLPAAGEFLDMITPQYLADL-MSWGAIGARTTESQVHR-ELASGLSCPVGFKNGTDGTIKVAIDAINAAGAPHCFLSVTKWG 217
Iβ
Bsu AroQ•AroA 169 ----VEGLQILKRVADEFDLAVISEIVTPAHIEEALDY-IDVIQIGARNMQNFELL-KAAGAVKKPVLLKRGLAATISEFINAAEYIMSQ---------- 252
Aae_KdsA 61 ----EYGVKALRKVKEEFGLKITTDIHESWQAEPVAEV-ADIIQIPAFLCRQTDLL-LAAAKTGRAVNVKKGQFLAPWDTKN-VVEKLKFG--------- 144
Eco KdsA 74 ----EEGMKIFQELKQTFGVKIITDVHEPSQAQPVADV-VDVIQLPAFLARQTDLV-EAMAKTGAVINVKKPQFVSPGQMGNIVDKFKEG--------- 158
Iα
Eco_AroA_F 218 HSAIVNTSGNGDCHIILRGG-K-----EPNYSAKHVAEVKEGLNKAGLPAQ------VMIDFSHAN----------SSKQFKKQMDVCADVCQQIAGGE 294
Iβ
Bsu AroQ•AroA 253 --------GNDQIILCERGIRTYET---ATRNTLDISAVPILKQET-H-------LPVFVDVTH----------S---TGRRDLLLPTAKAALAIG- 316
Aae_KdsA 145 --------GAKEIYLTERGTTFGYN-----NLVVDFRSLPIMKQW----------AKVIYDATHSVQLPGGL-----GDKSGGMREFIFPLIRAAVAVG- 215
Eco KdsA 159 --------GNEKVILCDRGANFGYD-----NLVVDMLGFSIMKKVSGN-------SPVIFDVTHALQCRDPF-----GAASGGRRAQVAELARAGMAVG- 232
Iα
Eco_AroA_F 295 KA---IIGVMVESHLVEGNQSLESGEP-------LAYGKSITDACIGWEDTDALLRQLANAVKARRG 351
Iβ
Bsu AroQ•AroA 317 -----ADGVMAEVHPDPSVALS--------------------DSAQQMAIPEFEKWLNELKPMVKVNA 359
Aae KdsA 216 -----CDGVFMETHPEPEKALS--------------------DASTQLPLSQLEGIIEAILEIREVASKYYETIPVK 267
Eco KdsA 233 -----LAGLFIEAHPDPEHAKC--------------------DGPSALPLAKLEPFLKQMKAIDDLVKGFEELDTSK 284

emplified by the *Bacillus subtilis* enzyme). With the limited number of sequences available at that time, $AroA_{I\beta}$ appeared to constitute a relatively small subfamily. However, it is now apparent that $AroA_{I\beta}$ is in fact the most widely distributed type of DAHP synthase, and therefore the ancestral form of AroA was most likely the Iβ type (Gosset et al. 2001). KDOP synthases (KdsA) were found to join $AroA_{I\beta}$ DAHP synthases as a second group within the I_β clade of proteins (Subramaniam et al. 1998). Thus, the I_β proteins with AroA activity exhibit more overall similarity to the proteins having KdsA activity than they do to the I_α proteins having AroA activity. In this case, the substrate specificity that defines DAHP synthase sorts into clusters at two hierarchical levels. It would be appropriate to name the Class I proteins the 3-deoxy-ald-2-ulosonate-phosphate synthase family to accommodate the different substrate specificities of the member subfamilies.

The specific phylogenetic relationship of KdsA with $AroA_{I\beta}$, as well as the relationship of both with $AroA_{I\alpha}$ at a deeper level was deduced from detailed analyses of multiple sequence alignments. Alignments between $AroA_{I\alpha}$ and $AroA_{I\beta}$ and between $AroA_{I\beta}$ and KdsA were carried out to locate conserved motifs that reflect the common substrate specificities of $AroA_{I\alpha}$ and $AroA_{I\beta}$, on the one hand, and between $AroA_{I\beta}$ and KdsA, on the other hand, that reflect their relatively recent ancestry. In addition, variable N-terminal regions that encode a variety of independently evolved allosteric regions for DAHP synthase with no counterparts in KdsA were excised to prevent these from pulling the alignment of homologous residues in the basic catalytic core region out of register. These alignments were then used to make extensive manual adjustments to the overall alignment. While no tree analysis was done, subsequent structural information (Shumilin et al. 1999; Radaev et al. 2000; Wagner et al. 2000a, b; Duewel et al. 2001; Asojo et al. 2001) has reinforced the overall validity of the alignment. Furthermore, new publicly available tools that have gained wide acceptance for analysis of remote relationships affirmed the homology relationship between KdsA and all AroA members of Class I. Thus, the use of the PSI-Blast tool at NCBI (www.ncbi.nlm.nih.gov) returns significant pairwise alignments for KdsA, $AroA_{I\alpha}$, and $AroA_{I\beta}$ after relatively few iterations, using given KdsA, $AroA_{I\alpha}$, and $AroA_{I\beta}$ proteins as alternative queries. Birck and Woodard (2001) belatedly hypothesize a common ancestry of groupings that has been apparent for some time.

With the additional sequences now available in the database and with further guidance from recent structural information, a new and updated multiple alignment has been generated (available upon request). Initial alignments were generated by input of almost 100 sequences (identified in Table 1) into the ClustalW program (version 1.4). Manual alignment adjustments were made to bring conserved motifs and residues into register. This procedure was facilitated by use of the BioEdit multiple alignment tool (Hall 2001). An additional refinement was the generation of a multiple alignment in which any N-terminal extensions lacking in kdsA proteins were excised. This produces an alignment corresponding to the core catalytic region. Both the refined and the whole-sequence multiple alignments were used as input for phylogenetic tree analysis using the program packages PHYLIP (Felsenstein 1989) and PHYLO_WIN (Galtier et al. 1996). The Protpars program was used to generate a maximum parsimony tree, and the neighbor-joining and Fitch programs were employed to obtain distance-based trees. The distance matrix used for the latter programs was obtained using Protdist with a Dayhoff PAM matrix. The Seqboot and Consense programs were then used to assess the statistical strength of the tree using bootstrap resampling. Neighbor-joining and Fitch distance trees and maximum parsimony trees resulted in similar clusters and arrangement of taxa within them.

Figure 1 displays a convenient and compact template for visualization of conserved motifs defined by invariant residues within the Family I 3-deoxy-ald-2-ulosonate-phosphate synthases. Invariant residues are displayed by color coding at the hierarchical levels of $AroA_{I\alpha}$, $AroA_{I\beta}$, and KdsA. The distinctive *Streptococcus* sequences occupy an outlying position in the $AroA_I$ clade (see Fig. 2), and Fig. 1 shows in orange the positions of 10 residues for which *Streptococcus* is the only exception to invariance in family $AroA_{I\alpha}$. Vertical bars join invariant residues that are common to both $AroA_{I\alpha}$ and $AroA_{I\beta}$ (5 residues), to both $AroA_{I\beta}$ and KdsA (4 residues), or to the entire family (11 residues). Many other matches can be found that approach invariance, especially near the invariant residues. Invariant residues clearly must have important functional roles. Three of

Fig. 1. Pattern of invariant residues conserved at different hierarchical levels within the Family I 3-deoxy-ald-2-ulosonate-phosphate synthases. The alignment positions of Eco $AroA_F$ (representing 46 $AroA_{I\alpha}$ members), the AroA domain of Bsu AroQ·AroA (representing 19 $AroA_{I\beta}$ members), and Aae KdsA and Eco KdsA (representing 31 KdsA members) were extracted from a multiple alignment (available upon request). The gaps required in the overall alignment are shown as *dashes.* Eco $AroA_F$, Eco KdsA, and Aae KdsA were studied by X-ray crystallography. The AroA proteins from *Streptococcus* are a distinctive divergent lineage within the Iα subfamily (see Fig. 2). Amino acid residues of *Streptococcus* $AroA_{I\alpha}$ proteins are highlighted in *orange* (second line) to show the 10 residues that are otherwise invariant in subfamily Iα, differing only in *Streptococcus.* Residue numbers are shown at the *left* and the *right.* Residues whose invariance is restricted to the $AroA_{I\alpha}$, $AroA_{I\beta}$, or KdsA grouping are highlighted in *yellow,* and positions exhibiting only a single exception are displayed in *gray. Vertical bars* connect residues conserved between $AroA_{I\alpha}$ and $AroA_{I\beta}$ (*blue*), between $AroA_{I\beta}$ and KdsA (*pink*), or between the entire protein family (*green*). Metal-coordinating residues are labeled M^{++}.

Table 1. Compilation of gene identification (gi) numbers of sequences used[a]

Organism[b]	Subfamily Iα AroA	Subfamily Iβ AroA	Subfamily Iβ KdsA
Actinobacillus actinomycetemcomitans	n.a		
Actinobacillus pleuropneumoniae			gi 6647547
Aeropyrum pernix		gi 7436657	
Amycolatopsis mediterranei	gi 13492244		
Amycolatopsis orientalis	gi 7522110		
Aquifex aeolicus			gi 6647545
Arabidopsis thaliana			gi 4835760
Bacillus halodurans		gi 10175865	
Bacillus subtilis		gi 728897	
Bordetella bronchiseptica	n.a.		n.a.
Bordetella bronchiseptica	n.a.		
Bordetella pertussis	n.a.		n.a.
Bordetella pertussis	n.a.		
Buchnera aphidicola	gi 1168514 ($AroA_W$)		
Campylobacter jejuni			gi 6967858
Candida albicans	gi 461538 ($AroA_F$)		
Candida albicans	gi 2492964 ($AroA_Y$)		
Caulobacter crescentus	gi 13422754		gi 3422792
Chlamydia muridarum		gi 7190696	gi 7190065
Chlamydia pneumoniae		gi 7436658	gi 6647568
Chlamydia trachomatis		gi 7436659	gi 1644362
Chlamydia psittaci			gi 2498497
Chlorobium tepidum	n.a.		n.a.
Clostridium acetobutylicum		n.a.	
Clostridium difficile		n.a.	
Clostridium difficile		n.a.	
Corynebacterium glutamicum	gi 461539		
Deinococcus radiodurans	gi 7473312	gi 7473313	
Enterococcus faecalis		n.a.	
Escherichia coli	gi 114195 ($AroA_F$)		
Escherichia coli	gi 114196 ($AroA_W$)		
Escherichia coli	gi 114190 ($AroA_Y$)		gi 1708631
Erwinia herbicola	gi 6225069 ($AroA_W$)		
Haemophilus influenzae	gi 1168513		gi 1170637
Helicobacter pylori			gi 2498499
Klebsiella pneumoniae	n.a. ($AroA_F$)		
Mesorhizobium loti			gi 13470612
Mycobacterium avium	n.a.		
Neisseria gonorrhoeae	n.a.		n.a.
Neisseria meningitides	gi 7225527		gi 7226523
Pasteurella haemolytica			gi 2498500
Pasteurella multocida	n.a.		n.a.
Pasteurella multocida	n.a.		
Pisum sativum			gi 6647535
Porphyromonas gingivalis			n.a.
Pseudomonas aeruginosa	gi 11350767		gi 6647573
Pseudomonas aeruginosa	gi 11350768		
Pyrococcus abyssi		gi 7436656	
Pyrococcus furiosus		n.a.	
Rickettsia prowazekii			gi 6647572
Saccharomyces cerevisiae	gi 6320240 ($AroA_F$)		
Saccharomyces cerevisiae	gi 6319726 ($AroA_Y$)		
Salmonella typhimurium	n.a. ($AroA_F$)		n.a.
Salmonella typhimurium	n.a. ($AroA_W$)		
Salmonella typhimurium	gi 114191 ($AroA_Y$)		
Schizosaccharomyces pombe	gi 1168512		
Shewanella putrefaciens	n.a.		n.a.
Shewanella putrefaciens	n.a.		
Shewanella putrefaciens	n.a		
Staphylococcus aureus		gi 13701532	
Staphylococcus xylosus		gi 1177684	
Streptococcus mutans	n.a.		
Streptococcus mutans	n.a.		
Streptococcus pneumoniae	n.a.		

Table 1. Continued

Organism[b]	Subfamily Iα AroA	Subfamily Iβ AroA	Subfamily Iβ KdsA
Streptococcus pneumoniae	n.a.		
Streptococcus pyogenes		n.a.	
Synechococcus sp.		gi 6012178	
Thermoplasma acidophilum		gi 10639390	
Thermoplasma volcanium		gi 13542182	
Thermotoga maritima		gi 7448834	
Thiobacillus ferrooxidans		n.a.	n.a.
Vibrio cholerae	gi 8118490		gi 11268173
Vibrio cholerae	gi 9655135		
Vibrio cholerae	VCA1036 (TIGR)		
Yersinia pestis	n.a. ($AroA_F$)		
Yersinia pestis	n.a. ($AroA_W$)		
Yersinia pestis	n.a. ($AroA_Y$)		
Xylella fastidiosa			gi 11268172

[a] For those sequences of unfinished genomes where gi numbers are not available (n.a.), FASTA files are available from the authors.
[b] The corresponding acronym identifiers used in Fig. 1 comprise the first letter of the genes and the first two letters of the species, e.g., *Aeropyrum pernix* is Ape.

the four established metal ligands are invariant. C-11 (Aae KdsA numbering) is conserved in the majority of KdsA sequences (see Fig. 2) but is substituted in the remainder (as illustrated by N-26 of Eco KdsA). The functional roles of some invariant residues for coordination with substrates have been elucidated by the various studies of the crystal structures, and ongoing work should be additionally informative.

Figure 2 shows results obtained from the neighbor-joining program displayed as an unrooted tree. The AroA proteins split into two groups consistent with the $Aroa_{I\alpha}$ and $AroA_{I\beta}$ divisions asserted by Subramaniam et al. (1998) and similar to the maximum likelihood tree presented (their Fig. 3) by Birck and Woodard (2001), except for their unacceptable inclusion of two $AroA_{II}$ sequences. Figure 2 also shows the greater proximity of $AroA_{I\beta}$ to KdsA than to $AroA_{I\alpha}$. Although Birck and Woodard (2001) depict the KdsA proteins as a division of two distinct groups (their Fig. 2), there is in fact no divergence that is at all comparable to the divergence between $AroA_{I\alpha}$ and $AroA_{I\beta}$. Rather, their "Class I" group (shown in yellow) is a normal divergence that generally parallels expectations for 16S rRNA trees of the γ-division Proteobacteria themselves. The much heavier representation of sequences in the "Class I" lineage than in any of the other lineages at an equivalent hierarchical level can give a visually misleading impression of bifurcation.

Birck and Woodard (2001) have speculated that $AroA_{I\alpha}$ proteins are metalloenzymes and $AroA_{I\beta}$ proteins are not. This may very well be, but the sole support thus far is the observation that the *Bacillus subtilis* enzyme is not sensitive to EDTA treatment (first shown by Jensen and Nester 1966). Note that the *B. subtilis* enzyme, like all other $AroA_{I\beta}$ proteins, has all four metal-liganding residues (unlike Eco KdsA). Thus, if it proves to be correct that $AroA_{I\beta}$ proteins are nonmetalloproteins, it is not just a simple matter of whether the initial cysteine metal-liganding residue is present or not. Although the separation of KdsA proteins into the metallo- and nonmetalloprotein grouping proposed (Birck and Woodard 2001) requires much more documentation, a rigorous example of each type does exist in support of this possibility.

In Fig. 2 the "Class I" DAHP synthases ($AroA_{I\beta}$) and the "Class I" KdsA proteins, defined by Birck and Woodard (2001) as enzymes lacking any metal requirement, are shown in yellow. Recently, an additional example of a nonmetallo KdsA protein has been described (Brabetz et al. 2000) that falls outside the "Class I" grouping (highlighted in orange in Fig. 2). Birck and Woodard (2001) dismiss the validity of obtaining any tree relationship of KDOP synthase and DAHP synthase on the grounds of "low sequence similarity." Yet the $AroA_{I\alpha}$ and $AroA_{I\beta}$ groups are included together in one of their trees, despite the fact that these are more remote from one another than are $AroA_{I\beta}$ and KdsA (Fig. 2). They imply that "Class II" KDOP synthase and "Class II" DAHP synthases (defined as metalloenzymes) group together based upon the use of common metal-coordinating residues. Likewise, evolutionary linkage of "Class I" KDOP synthase and "Class I" DAHP synthase was asserted, based primarily on the alteration of a single metal-coordinating residue. They therefore assert an evolutionary relationship between the KdsA and the $AroA_{I\beta}$ proteins shown in yellow in Fig. 2, on the one hand, and between $AroA_{I\alpha}$ and the remaining KdsA proteins, on the other hand. However, Fig. 2 shows that KdsA proteins, whether they be metalloenzymes or nonmetalloenzymes, are more closely related to one another than to

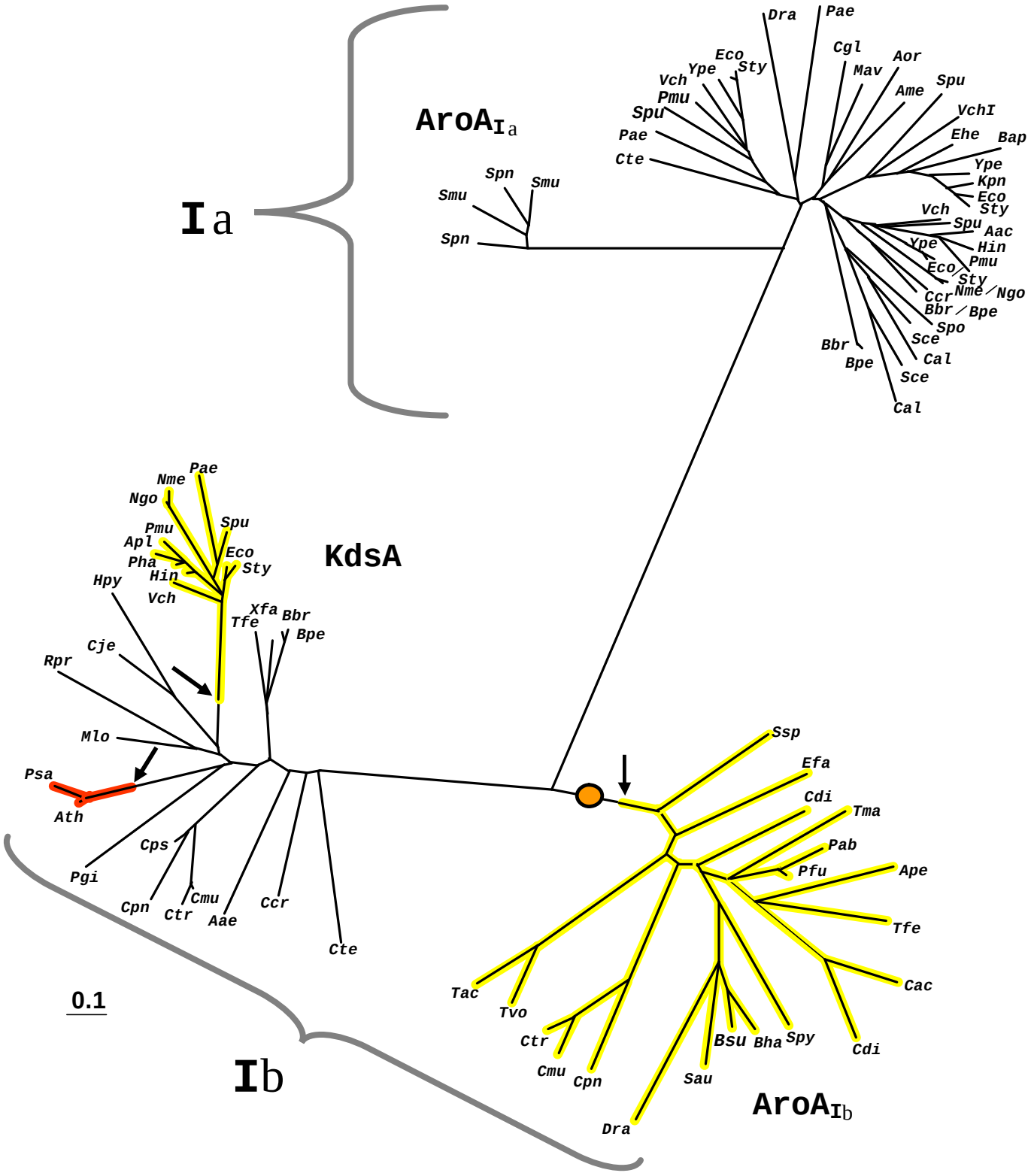


Fig. 2. Unrooted phylogenetic tree (radial view) of homology Class I proteins consisting of Family Ia AroA and Family Iß AroA and KdsA subfamilies. The tree was generated with the neighbor-joining method. Yellow highlighting shows the "Class I" KDOP synthases on the lower left and the "Class I" DAHP synthases on the lower right (our AroA $_{Iß}$) that were asserted by Birck and Woodward (2001). The lineage leading to two higher plant KdsA sequences is highlighted in orange. A hypothetical root is indicated with a circle, and separate evolutionary events of loss of dependence upon metal for catalysis that are postulated are shown with arrows. Bootstrap values of 1000 per 1000 iterations supported the major nodes, and the order of branching within the various component clusters was generally supported with high bootstrap values. The manually adjusted multiple alignment used as input for the tree program contained complete sequences in order to better relate the results to those of Birck and Woodward (2001). The branching relationships within each major cluster are slightly different when extraneous N-terminal extensions were excised in order to analyze only the core catalytic region. The latter tree and input multiple alignment are available upon request.

any AroA protein, whether metalloprotein ($AroA_{I\alpha}$) or putative nonmetalloprotein ($AroA_{I\beta}$). At the higher hierarchical level KdsA and $AroA_{I\beta}$ are more similar to one another than to $AroA_{I\alpha}$, regardless of the presence or absence of a metal requirement.

Although the tree shown in Fig. 2 is an unrooted tree, a reasonable speculation would be to place the root of the tree nearest the $AroA_{I\beta}$ grouping since members of this group are the most widely distributed in nature (Gosset et al. 2001). A suggested evolutionary scenario (Fig. 2) that accommodates the metalloprotein character state with a correct tree relationship requires at least three independent events where metal dependence for catalysis was lost. The assumption of the ancient existence of metal-coordinating residues as a primitive catalytic scaffold seems in line with current thinking. Three separate events of loss are postulated: in the common ancestor of the entire $AroA_{I\beta}$ group, within the KdsA group in the lineage of divergence to the γ division of Proteobacteria, and within the KdsA group in the lineage leading to higher plants. Since loss of metal dependence may happen easily, it would not be surprising to see additional independent evolutionary events of loss of metal dependence as more sequences become available.

The recruitment hypothesis (Jensen 1976) for acquisitive evolution and expansion of the metabolic repertoire asserts that ancient proteins generally had broad substrate specificities prior to events of gene duplication that facilitated differential narrowing of substrate specificities and regulation. In this context it was intriguing to find a KDOP synthase in *Neisseria gonorrhoeae* that could utilize (weakly) erythrose-4-P (E4P) or ribose-5-P in place of arabinose-5-P (Subramaniam et al. 1998). A very similar broad-specificity KDOP synthase was found in higher plants (Doong et al. 1991; Doong and Jensen 1992). This broad substrate specificity was not found in recombinant KdsA from *N. gonorrhoeae* by Sheflyan et al. (2000). Among the possible reasons for this discrepancy are the following. (i) Technical differences in enzyme assay might exist, e.g., E4P dimerizes or dephosphorylates under some conditions and is then inactive as a substrate. E4P is notorious for promoting substrate inhibition, and the 33 m*M* concentrations used by Sheflyan et al. (2000) exceed ours by well over an order of magnitude. E4P may also contain contaminating sugars that may be inhibitory. (ii) Recombinant protein might be produced under different conditions of the intracellular microenvironment that affect the properties of the protein. As one would expect from an alternative substrate, we found that E4P was a competitive inhibitor of A5P utilization by *N. gonorrhoeae* KDOP synthase (Subramaniam et al. 1998). Although Sheflyan et al. (2000) did not find activity with E4P, it has been concluded that E4P binds in the active site since KDOP synthase utilization of A5P was inhibited by E4P in *E. coli* (Duewel et al. 2001). This indication that E4P can access the active site might mean that particular conditions of the microenvironment in the normal host cell might be required for activity with E4P. (iii) Perhaps a different gene product was assayed by Subramaniam et al. (1998) and by Sheflyan et al. (2000). We have recently found that *N. gonorrhoeae* (n.a.) and *N. meningitides* (gi 11352947) have *aro*$A_{I\beta}$ genes in addition to the single *aro*$A_{I\alpha}$ genes present (Gosset et al. 2001). Although these encoded proteins are clear $AroA_{I\beta}$ homologues (e.g., 30% identity of Ngo $AroA_{I\beta}$ with Tfe $AroA_{I\beta}$; expect value = 4*e*-30 based upon Blast query of the nr database), many of the conserved residues that are crucial for DAHP synthase activity (Shumilin et al. 1999) have been altered. Perhaps this apparent pseudogene encodes the broad-specificity KDOP synthase that was studied by Subramaniam et al. (1998). If so, the PEP atomic contacts that dictate stereospecificity (Asojo et al. 2001) may have been completely remodeled.

Although so stated by Birck and Woodard (2001), the latter issue of substrate specificity, which remains to be resolved, has no bearing upon the validity of the original assertion of Subramaniam et al. (1998) that all KdsA proteins belong to a common phylogenetic grouping that includes $AroA_{I\beta}$.

Acknowledgment. G.X. was partially supported in this work through the STDGEN project at Los Alamos National Laboratory (NIH/NIAID IAG Y1-A1-8228-05).

References

Asojo O, Friedman J, Adir N, Belakhov V, Shoham Y, Baasov T (2001) Crystal structures of KDOP synthase in its binary complexes with the substrate phosphoenolpyruvate and with a mechanism-based inhibitor. Biochemistry 40:6326–6334

Birck MR, Woodard RW (2001) *Aquifex aeolicus* 3-deoxy-*D-manno*-2-octulosonic acid 8-phosphate synthase: A new class of KDO 8-P synthase? J Mol Evol 52:205–214

Brabetz W, Wolter FP, Brade H (2000) A cDNA encoding 3-deoxy-*D-manno*-oct-2-ulosonate-8-phosphate synthase of *Pisum sativum* L. (pea) functionally complements a *kds*A mutant of the Gram-negative bacterium *Salmonella enterica.* Planta 212:136–143

Doong RL, Jensen RA (1992) Synonymy of the three apparent isoenzymes of 3-deoxy-*D-arabino*-heptulosonate 7-phosphate synthase in *Pisum sativum* L. with 3-deoxy-*D-manno*-octulosonate 8-phosphate synthase and the DS-Co/DS-Mn isoenzyme pair. New Phytol 121:165–171

Doong RL, Ahmad S, Jensen RA (1991) Higher plants express 3-deoxy-*D-manno*-octulosonate 8-phosphate synthase. Plant Cell Environ 14:113–120

Duewel HS, Radaev S, Wang J, Woodard RW, Gatti DL (2001) Substrate and metal complexes of 3-deoxy-*D-manno*-octulosonate-8-phosphate synthase from *Aquifex aeolicus* at 1.9-Å resolution. J Biol Chem 276:8393–8402

Felsenstein J (1989) PHYLIP—Phylogeny inference package (version 3.2). Cladistics 5:164–166

Galtier N, Gouy M, Gautier C (1996) SEAVIEW and PHYLO_WIN:

Two graphic tools for sequence alignment and molecular phylogeny. Comput Appl Biosci 12(6):543–548

Gosset G, Bonner CA, Jensen RA (2001) Microbial origin of plant-type 2-keto-3-deoxy-*D-arabino*-heptulosonate 7-phosphate synthases, exemplified by the chorismate and tryptophan-regulated enzyme from *Xanthomonas campestris.* J Bacteriol 183:4061–4070

Hall T (2001) Biological sequence alignment editor for Windows 95/98/NT, 5.09 ed. North Carolina State University, Raleigh (www.mbio.ncsu.edu/BioEdit/bioedit.html)

Jensen RA (1976) Enzyme recruitment in evolution of new function. Annu Rev Microbiol 30:409–425

Jensen RA, Nester EW (1966) Regulatory enzymes of aromatic amino acid biosynthesis in *Bacillus subtilis.* J Biol Chem 241:3365–3372

Radaev S, Dastidar P, Patel M, Woodard RW, Gatti DL (2000) Structure and mechanism of 3-deoxy-*D-manno*-oculosonate 8-phosphate synthase. J Biol Chem 275:9476–9484

Sheflyan GY, Sundaram AK, Taylor WP, Woodard RW (2000) Substrate ambiguity of 3-deoxy-*D-manno*-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* revisited. J Bacteriol 182: 5005–5008

Shumilin IA, Kretsinger RH, Bauerle RH (1999) Crystal structure of phenylalanine-regulated 3-deoxy-*D-arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli.* Struct Fold Des 7:865–875

Subramaniam PS, Xie G, Xia T, Jensen RA (1998) Substrate ambiguity of 3-deoxy-*D-manno*-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* in the context of its membership in a protein family containing a subset of 3-deoxy-*D-arabino*-heptulosonate 7-phosphate synthases. J Bacteriol 180:119–127

Wagner T, Kretsinger RH, Bauerle R, Tolbert WD (2000a) 3-Deoxy-*D-manno*-octulosonate-8-phosphate synthase from *Escherichia coli.* Model of binding of phosphoenolpyruvate and *D*-arabinose-5-phosphate. J Mol Biol 301:233–238

Wagner T, Shumilin IA, Bauerle R, Kretsinger RH (2000b) Structure of 3-deoxy-*D-arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli:* Comparison of the Mn^{2+} *2-phosphoglycolate and the Pb^{2+} *2-phospho*enol*pyruvate complexes and implications for catalysis. J Mol Biol 301:389–399

Walker GE, Dunbar B, Hunter IS, Nimmo HG, Coggins JR (1996) Evidence for a novel class of microbial 3-deoxy-*D-arabino*-heptulosonate-7-phosphate synthase in *Streptomyces coelicolor* A3(2), *Streptomyces rimosus* and *Neurospora crassa.* Microbiology 142:1973–1982