# Substrate Ambiguity and Crystal Structure of *Pyrococcus furiosus* 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate Synthase: An Ancestral 3-Deoxyald-2-ulosonate-phosphate Synthase?[†,‡]

Linley R. Schofield,[§] Bryan F. Anderson,[||] Mark L. Patchett,[||] Gillian E. Norris,[||] Geoffrey B. Jameson,[§] and Emily J. Parker*,[§]

*Institutes of Fundamental Sciences and Molecular BioSciences, Massey University, Palmerston North, New Zealand*

*Received March 29, 2005; Revised Manuscript Received May 22, 2005*

ABSTRACT: 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (DAH7PS) catalyzes the condensation reaction between phosphoenolpyruvate (PEP) and the four-carbon monosaccharide D-erythrose 4-phosphate (E4P). DAH7PS from the hyperthermophile *Pyrococcus furiosus* is a member of the DAH7PS I$\beta$ subfamily, which also includes the KDO8PS enzymes. KDO8PS (3-deoxy-D-*manno*-octulosonate-8-phosphate synthase) catalyzes a closely related reaction of PEP with the five-carbon monosaccharide D-arabinose 5-phosphate (A5P). DAH7PS from *P. furiosus* requires a metal ion for activity and, unlike other characterized DAH7PS enzymes, is not inhibited by aromatic amino acids. Purified *P. furiosus* DAH7PS is able to utilize not only the four-carbon phosphorylated monosaccharides E4P and 2-deoxy-D-erythrose 4-phosphate but also the five-carbon phosphorylated monosaccharides A5P, D-ribose 5-phosphate, and 2-deoxy-D-ribose 5-phosphate with similar $k_{cat}$ but much increased $K_M$ values. DL-Glyceraldehyde 3-phosphate and D-glucose 6-phosphate are not substrates. The structure of recombinant *P. furiosus* DAH7PS in complex with PEP was determined to 2.25 Å resolution. The asymmetric unit consists of a dimer of $(\beta/\alpha)_8$-barrel subunits. Analysis of the buried surfaces formed by dimerization and tetramerization, as observed in the crystal structure, provides insight into both the oligomeric status in solution and the substrate ambiguity of *P. furiosus* DAH7PS. *P. furiosus* DAH7PS is both the first archaeal and the first "naked" DAH7PS (without N-terminal extensions) to be fully characterized functionally and structurally. The broad substrate specificity of this DAH7PS, the lack of allosteric inhibition, and various structural features indicate that, of the enzymes characterized to date, *P. furiosus* DAH7PS may be the contemporary protein closest to the ancestral type I enzyme.

The enzyme 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (DAH7PS,[1] EC 2.5.1.54) catalyzes the condensation of phosphoenolpyruvate (PEP) and a four-carbon phosphorylated sugar D-erythrose 4-phosphate (E4P) to form 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAH7P) (Figure 1). This reaction is the first step in the shikimate pathway that is used to synthesize chorismate, the precursor of the aromatic amino acids phenylalanine, tyrosine, and tryptophan, and of many aromatic secondary metabolites (*1*). The shikimate pathway is found in plants and microorganisms, but not in animals (*2*). In a closely related reaction, 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase (KDO8PS, EC 2.5.1.55) catalyzes the condensation of PEP and a five-carbon phosphorylated sugar D-arabinose 5-phosphate (A5P) to form 3-deoxy-D-*manno*-octulosonate 8-phosphate (KDO8P) (Figure 1). This reaction is a key step in the biosynthesis of 3-deoxy-D-*manno*-octulosonate (KDO), which is required for lipopolysaccharide biosynthesis in Gram-negative bacteria (*3*). As neither DAH7PS nor KDO8PS has a known counterpart in mammalian metabolism (*4*), both enzymes are attractive targets for antimicrobial compounds.

Although DAH7PS and KDO8PS belong to functionally unrelated pathways, a common ancestor has been inferred. Both enzymes catalyze the condensation of PEP with a phosphorylated aldose by a similar ordered-sequential mechanism [PEP/E4P(A5P)//$PO_4^{3-}$/DAH7P(KDO8P)] involving the cleavage of the C−O bond of PEP rather than the more usual P−O bond cleavage (*5*, *6*). This requires reaction mechanisms in which the anomeric oxygen is derived from water. Additionally, both reactions are highly stereospecific with the *si* face of PEP coupling with the *re* face of their

[†] This work was supported by the Royal Society of New Zealand Marsden Fund (MAU008).

[‡] The coordinates of the refined structure have been deposited with the Protein Data Bank (entry 1ZCO).

* To whom correspondence should be addressed. Telephone: (+64) 6 350-5799, ext. 3566. Fax: (+64) 6 350-5682. E-mail: E.J.Parker@massey.ac.nz.

[§] Institute of Fundamental Sciences.

[||] Institute of Molecular BioSciences.

[1] Abbreviations: A5P, D-arabinose 5-phosphate; BTP, 1,3-bis[tris(hydroxymethyl)methylamino]propane; DAH7P, 3-deoxy-D-*arabino*-heptulosonate 7-phosphate; DAH7PS, 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase; 2dE4P, 2-deoxy-D-erythrose 4-phosphate; 2dR5P, 2-deoxy-D-ribose 5-phosphate; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; E4P, D-erythrose 4-phosphate; G6P, D-glucose 6-phosphate; G3P, DL-glyceraldehyde 3-phosphate; KDO, 3-deoxy-D-*manno*-octulosonate; KDO8P, 3-deoxy-D-*manno*-octulosonate 8-phosphate; KDO8PS, 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase; MWCO, molecular weight cutoff; PAGE, polyacrylamide gel electrophoresis; PEG, polyethylene glycol; PEP, phosphoenolpyruvate; R5P, D-ribose 5-phosphate; rms, root-mean-square.

10.1021/bi050577z CCC: $30.25 
Published on Web 08/16/2005

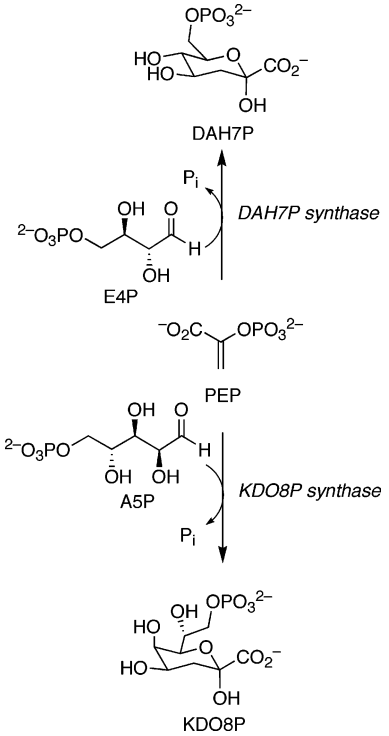



FIGURE 1: Condensation of PEP with phosphorylated monosaccharides.

respective monosaccharide substrates (*7*, *8*). Evidence for a common ancestor has been strengthened as additional sequence information has become available. Moreover, the X-ray crystal structures of DAH7PS from bacterial (*Escherichia coli* and *Thermotoga maritima*) and eukaryotic (*Saccharomyces cerevisiae*) organisms have been found to be remarkably similar to those of KDO8PS from bacterial sources (*Aquifex aeolicus* and *E. coli*) (*9*−*13*). The DAH7PS and KDO8PS structures all share a similar homodimeric arrangement of $(\beta/\alpha)_8$-barrel subunits, in which the active sites are in proximity and the active site residues are generally found in structurally equivalent positions. In the DAH7PS from *T. maritima* (*Tm*-DAH7PS) and the KDO8PS from *E. coli* (*Ec*-KDO8PS) and *A. aeolicus* (*Aa*-KDO8PS), the dimers further associate into similar tetramers, whereas in the DAH7PS from *E. coli* (*Ec*-DAH7PS) and *S. cerevisiae* (*Sc*-DAH7PS), tetramers are formed via association of the dimer in a different way.

Recently, the phylogenetic relationship between DAH7PS and KDO8PS has been explored (*14*−*16*). DAH7PS proteins can be divided into two distinct families by amino acid sequence homology and molecular mass (*17*). Type I DAH7PS enzymes are mainly found in prokaryotic and archaeal organisms, although some eukaryotic examples have been identified [e.g., those from *S. cerevisiae* (*10*, *18*, *19*) and *Neurospora crassa* (*20*−*22*)]. Type II DAH7PS proteins were originally identified in plants (*1*) and are now known to encompass a diverse set of microbial proteins of which the plant proteins form a subcluster (*23*).

Type I enzymes can be further divided into Iα and I$\beta$ subfamilies (*14*). Subfamily Iα appears to consist entirely of DAH7PS proteins and includes the structurally and functionally characterized enzymes of *E. coli* (*9*, *24*−*28*) and *S. cerevisiae* (*10*, *18*, *19*). All characterized Iα DAH7PS proteins are metalloenzymes and are sensitive to feedback inhibition by aromatic amino acids. To date, the substrate specificity of subfamily Iα enzymes has been examined only for the phenylalanine-sensitive isozyme from *E. coli* [*Ec*-DAH7PS(phe)]. In addition to its natural substrate, E4P, this enzyme also shows limited enzymic activity with a range of five-carbon monosaccharides, including A5P, 2-deoxy-D-ribose 5-phosphate (2dR5P), and D-ribose 5-phosphate (R5P) (*28*).

Subfamily I$\beta$ consists of both DAH7PS (subfamily I$\beta_D$) and KDO8PS (subfamily I$\beta_K$) proteins. The DAH7PS enzymes of subfamily I$\beta_D$ are more similar in overall sequence to KDO8PS enzymes of subfamily I$\beta_K$ than to the functionally equivalent Iα subfamily of DAH7PS (*14*). Subfamily I$\beta_D$ proteins are putative metalloenzymes, based on evidence from *Pyrococcus furiosus* DAH7PS (*Pf*-DAH7PS) (*29*) and *Tm*-DAH7PS (*30*). However, only the latter enzyme is subject to feedback inhibition. Moreover, *Tm*-DAH7PS is reported to show a narrow substrate specificity, utilizing E4P but not A5P or R5P as a substrate (*30*). Interestingly, subfamily I$\beta_K$ contains both metalloenzymes and non-metalloenzymes, the best characterized being *Aa*-KDO8PS (*31*, *32*) and *Ec*-KDO8PS (*3*), respectively. Subfamily I$\beta_K$ enzymes are not subject to feedback inhibition (*14*) and appear to be very substrate specific, utlilizing only A5P and not R5P or E4P (*3*, *33*, *34*). It has been proposed that type I proteins be renamed the 3-deoxyald-2-ulosonate-phosphate synthase family, to accommodate the different substrate specificities of the member subfamilies (*15*).

In their detailed analyzes, Jensen et al. proposed that the ancestor of type I enzymes had a broad substrate specificity, had the ability to coordinate metal, was not subject to allosteric inhibition, and was most likely the I$\beta$ type of DAH7PS, as this type is the most widely distributed in nature (*14*, *15*). Over time, a range of changes occurred to varying degrees: substrate specificity narrowed both in the stereospecificity of the reaction and in the specificity for the aldose, allosteric inhibition was acquired, and metal dependency was lost in some cases (*14*).

*P. furiosus* is a hyperthermophile, found in deep-sea hydrothermal vents and having an optimum growth temperature of 100 °C (*35*). The purification and characterization of *Pf*-DAH7PS have been previously described (*29*). This paper links the broad substrate specificity with the crystal structure of *Pf*-DAH7PS. The results presented here, together with the previous evidence of both the divalent metal ion requirement and the lack of feedback inhibition, indicate that *Pf*-DAH7PS may be the characterized contemporary protein closest to the ancestral type I DAH7PS, or 3-deoxyald-2-ulosonate-phosphate synthase. *P. furiosus* DAH7PS is both the first archaeal and the first "naked" DAH7PS (lacking N-terminal extensions) to be fully characterized, both functionally and structurally.

## EXPERIMENTAL PROCEDURES

*DAH7PS Activity Assay and Kinetic Studies.* The continuous assay of Schoner and Herrmann (*24*) was used and modified as previously described (*29*). The consumption of PEP was monitored at 232 nm [extinction coefficient at 232 nm of $2.6 \times 10^3$ $M^{-1}$ $cm^{-1}$ at 60 °C and pH 6.8 (*29*)]. One unit of activity is defined as the loss of 1 $\mu$mol of PEP per minute at the stated temperature. Assays to determine activity

with DL-glyceraldehyde 3-phosphate (G3P) and D-glucose 6-phosphate (G6P) included 400 $\mu$M PEP (Research Chemicals) and 16 mM G3P (Sigma) or 1 or 4 mM G6P (Sigma). Assays for the determination of kinetic parameters with E4P and 2-deoxy-D-erythrose 4-phosphate (2dE4P) were initiated by the addition of the four-carbon phosphorylated sugar. Assays with A5P, R5P, and 2dR5P were initiated by the addition of the purified *Pf*-DAH7PS. All kinetics experiments were performed in duplicate except those with A5P. To determine kinetic parameters for PEP, the reaction mixture contained either (i) 20−600 $\mu$M PEP with 480 $\mu$M E4P (Sigma), 66 $\mu$M 2dE4P [obtained by chemical synthesis from (*S*)-$\beta$-hydroxy-$\gamma$-butyrolactone (unpublished results)], 9 mM A5P (Sigma), or 4.5 mM R5P (Research Organics Inc.) or (ii) 28−280 $\mu$M PEP and 4.5 mM 2dR5P (Sigma). To determine kinetic parameters for phosphorylated monosaccharides, the reaction mixture contained 400 $\mu$M PEP with 5.2−104 $\mu$M E4P, 4.4−88 $\mu$M 2dE4P, 0.4−9.0 mM A5P, 0.2−4.0 mM R5P, or 0.2−4.0 mM 2dR5P.

*Protein Purification and Crystallization.* The recombinant *Pf*-DAH7PS used in these studies was isolated from *E. coli* BL21-CodonPlus(DE3)-RIL cells harboring plasmid pT7-PfuDAH7PS. The construction of pT7-PfuDAH7PS, growth of cultures, effective two-step purification, and characterization of *Pf*-DAH7PS have been described previously (*29*). The purified enzyme was washed with buffer [10 mM BTP containing 10 $\mu$M EDTA and 200 $\mu$M PEP at pH 6.8 and 25 °C, treated with Chelex-100 resin (Bio-Rad) and filtered], using a prewashed 2 mL Vivaspin 10 000 MWCO concentrator (Vivascience) to remove KCl and concentrate the enzyme.

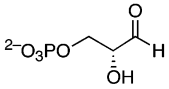
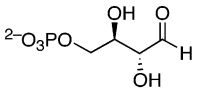
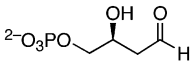
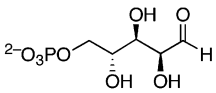
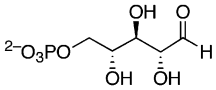
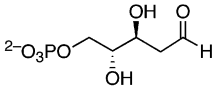
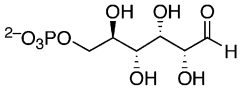
Crystals of *Pf*-DAH7PS were grown by vapor diffusion in a 2 $\mu$L hanging drop. For this procedure, the enzyme solution (9.2 mg/mL) was mixed 1:1 (v/v) with a reservoir solution containing 8% polyethylene glycol (PEG) 8000 (Hampton), 0.2 M ammonium acetate (Ajax Chemicals), and 0.1 M Tris-HCl (pH 7.3) (Hampton). The crystallization tray was incubated at 22 °C for 4 months. All solutions (including the *Pf*-DAH7PS solution) were filtered through a 0.2 $\mu$m membrane prior to use. The crystal was harvested, dipped in a cryoprotectant solution containing 4% PEG 8000, 0.1 M ammonium acetate, 0.05 M Tris-HCl (pH 7.3), and 50% glycerol, and frozen under a 110 K nitrogen stream.

*Data Measurement, Refinement, and Structural Analysis.* The structure of *Pf*-DAH7PS was determined by molecular replacement [MOLREP from the CCP4 suite (*36*)] from a hybrid model constructed from *Ec*-DAH7PS and *Aa*-KDO8PS. No extant KDO8PS or DAH7PS structure yielded molecular replacement solutions in a variety of programs. Given the relatively low level of sequence identity of *Pf*-DAH7PS to *Ec*-DAH7PS and to *Aa*-KDO8PS (26 and 30%, respectively), PsiPred (*37*) and mGenThreader (*38*) were used to predict secondary structure elements of *Pf*-DAH7PS and align these elements with those observed in structures of *Ec*-DAH7PS (PDB entry 1qr7) and *Ec*-KDO8PS (PDB entry 1d9e). The *Aa*-KDO8PS structure, which was not entered into the mGenThreader database, was structurally aligned with *Ec*-KDO8PS. SwissModel (*39*) was then used to generate coordinates for two possible models based upon the *Ec*-DAH7PS and *Aa*-KDO8PS structures. By reference to the mGenThreader results and observation of the motif for binding phosphorylated monosaccharides (R/K)xxxxK-PR(T/S), the SwissModel models were heavily edited and a hybrid model based upon the *Ec*-DAH7PS structure for residues 8−160 and the *Aa*-KDO8PS structure for residues 160−260 was constructed and subjected to molecular mechanics refinements within SwissModel. This refined model for *Pf*-DAH7PS had an rms difference for Cα positions of ∼0.4 Å with respect to the starting coordinates derived from the *Ec*-DAH7PS and *Aa*-KDO8PS structures. This difference, coupled with reasonable positions for many of the side chains, proved to be sufficient for determining the positions of the two nonequivalent molecules in the cell. The initial correlation coefficient was 0.23, but 20 rounds of phase extension with symmetry averaging (DM of the CCP4 suite) improved the correlation coefficient to a remarkable 0.70. The resulting electron density maps were easily interpretable, revealing a number of significant changes in several loops, including the $\beta$2−α2 loop, which in subunit B adopts a conformation significantly different from that seen in other DAH7PS structures, the α1−$\beta$2 loop, and the N- and C-terminal regions. There are also considerable differences in the interface region forming the canonical dimer. Despite EDTA treatment, residual electron density was found at the metal-binding site. The rms difference between the starting model and the final model is 0.81 Å for a superposition of 179 of 254 available Cα atoms, superimposing only those atoms meeting the criterion of lying within 1.5 Å of each other. Crystal parameters and data collection and refinement statistics are summarized in Table 1. Buried surface areas were calculated using Grasp (*40*) and are

Table 1: Crystal Parameters and Data Collection and Refinement Statistics for *Pf*-DAH7PS

| | |
|---|---|
| unit cell parameters | |
| crystal system | orthorhombic |
| space group | *I*222 |
| *a* (Å) | 87.23 |
| *b* (Å) | 110.02 |
| *c* (Å) | 144.35 |
| volume (Å$^3$) | 1385333 |
| *Z* (*Z*′) | 16 (2) |
| Matthews coefficient (solvent content) | 3.0 (58%) |
| data collection statistics | |
| (highest-resolution bin) | |
| no. of data collected | 166567 |
| no. of unique data | 33241 |
| resolution range (Å) | 19.80−2.25 (2.37−2.25) |
| redundancy | 4.97 (4.97) |
| completeness | 99.9 (100) |
| $R_{merge}$ | 0.126 (0.350) |
| $I/\sigma(I)$ | 7.6 (3.9) |
| refinement | |
| no. of protein residues | A, 1−262; B, 1−262 |
| no. of protein atoms | 8269 (4131 non-hydrogen) |
| no. of water molecules | 262 |
| other groups | 2 × PEP (0.5 occupancy) |
| | 1 × $M^{2+}$ (0.3 occupancy) |
| | 2 × $Cl^-$ (0.5 occupancy) |
| statistics (2.31−2.25 Å) | |
| *R* | 0.182 (0.276) |
| $R_{free}$ | 0.231 (0.338) |
| $R_{free}$ set | 2519, 7.6% |
| average *B* (protein) (Å$^2$) | 31.55 |
| average *B* (water) (Å$^2$) | 36.2 |
| rmsd for bond distances (Å) | 0.014 |
| rmsd for bond angles (deg) | 1.56 |
| Ramachandran parameters (%) | |
| most favored region | 92.8 |
| allowed region | 6.5 |
| generously allowed region | 0.4 |
| forbidden region | 0.2 ($N16_A$, $N16_B$) |
| PDB entry | 1ZCO |

Table 2: Kinetic Constants of *Pf*-DAH7PS with Different Monosaccharide Cosubstrates

| Substrate *structures* | $K_M$(monosaccharide) (μM) | $K_M$(PEP) (μM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_M$(monosacch.) ($\mu M^{-1} s^{-1}$) |
|---|---|---|---|---|
| G3P  | NS[a] | | | |
| E4P  | 9 ± 1 | 93 ± 9 | 1.4 ± 0.1 | 160 x $10^{-3}$ |
| 2dE4P  | 6 ± 1 | 33 ± 4 | 3.0 ± 0.1 | 490 x $10^{-3}$ |
| A5P  | 2700 ± 200 | 62 ± 8 | 1.1 ± 0.1 | 0.40 x $10^{-3}$ |
| R5P  | 1580 ± 110 | 36 ± 2 | 2.5 ± 0.1 | 1.6 x $10^{-3}$ |
| 2dR5P  | 2500 ± 150 | 35 ± 3 | 1.7 ± 0.1 | 0.69 x $10^{-3}$ |
| G6P  | NS | | | |

[a] NS, not a substrate.

provided in a comprehensive table found in the Supporting Information. Figures were prepared with Pymol (*41*).

## RESULTS

*Substrate Specificity and Kinetics.* A number of phosphorylated monosaccharides were tested as substrates, and kinetic parameters were determined for those showing activity (Table 2). No enzymic activity was detected using G3P or G6P, whereas all four- and five-carbon phosphorylated monosaccharides that were tested were found to be substrates. The five-carbon phosphorylated monosaccharides were considerably poorer substrates than their four-carbon counterparts, as a consequence of their relatively large $K_M$ values (but see Discussion). Although the values of the kinetic data for the five-carbon phosphorylated monosaccharides were similar, R5P was the better substrate, as evidenced by the highest $k_{cat}$ and $k_{cat}/K_M$ values, and the lowest $K_M$ value, compared to those of 2dR5P (or 2dA5P), which was similar to A5P. The most interesting result was the ability of 2dE4P to act not just as a substrate but as the most efficiently utilized substrate of those examined for *Pf*-DAH7PS, having a higher $k_{cat}$ and $k_{cat}/K_M$ and a lower $K_M$ than the in vivo substrate E4P.

*Crystal Structure Determination.* The crystal structure of *Pf*-DAH7PS was determined using molecular replacement methods and was refined at 2.25 Å resolution to a final *R* factor of 0.182 ($R_{free}$ = 0.231). The asymmetric unit of the crystal contains two molecules, and a 2-fold symmetry operation gives rise to a tetramer.

*Subunit Structure.* The subunit of *Pf*-DAH7PS consists of a $(\beta/\alpha)_8$-barrel domain, as illustrated in Figure 2. Figure 2a, a view looking down into the barrel, is rainbow-colored to highlight the progression of $\beta/\alpha$ motifs around the barrel. Each $\beta/\alpha$ motif, except for the $\beta3-\alpha3$ motif (colored aqua), contains at least one residue that contributes directly to the active site to which the metal, PEP, and E4P bind. Figure 2b gives a side-on view of the structure, in which the core structure common to all type I DAH7PS and KDO8PS structures characterized to date is highlighted. The noncore structure (i.e., not shared by all type I DAH7PS and KDO8PS enzymes) is colored green, and side chains are colored according to binding role. All of the substrate- and metal-binding residues come from the C-terminal ends of $\beta$-strands or the $\beta n-\alpha n$ loops. Long loops extend from the C-terminal ends of $\beta$-strands, especially those $\beta n-\alpha n$ loops for which $n$ = 2, 6, and 8, each of which is 15 residues long. On the other hand, the loops at the C-terminal ends of the α-helices are quite short, being one to three residues long (see Figure 2b). The bottom of the barrel is closed by the N-terminal two-stranded $\beta$-hairpin.

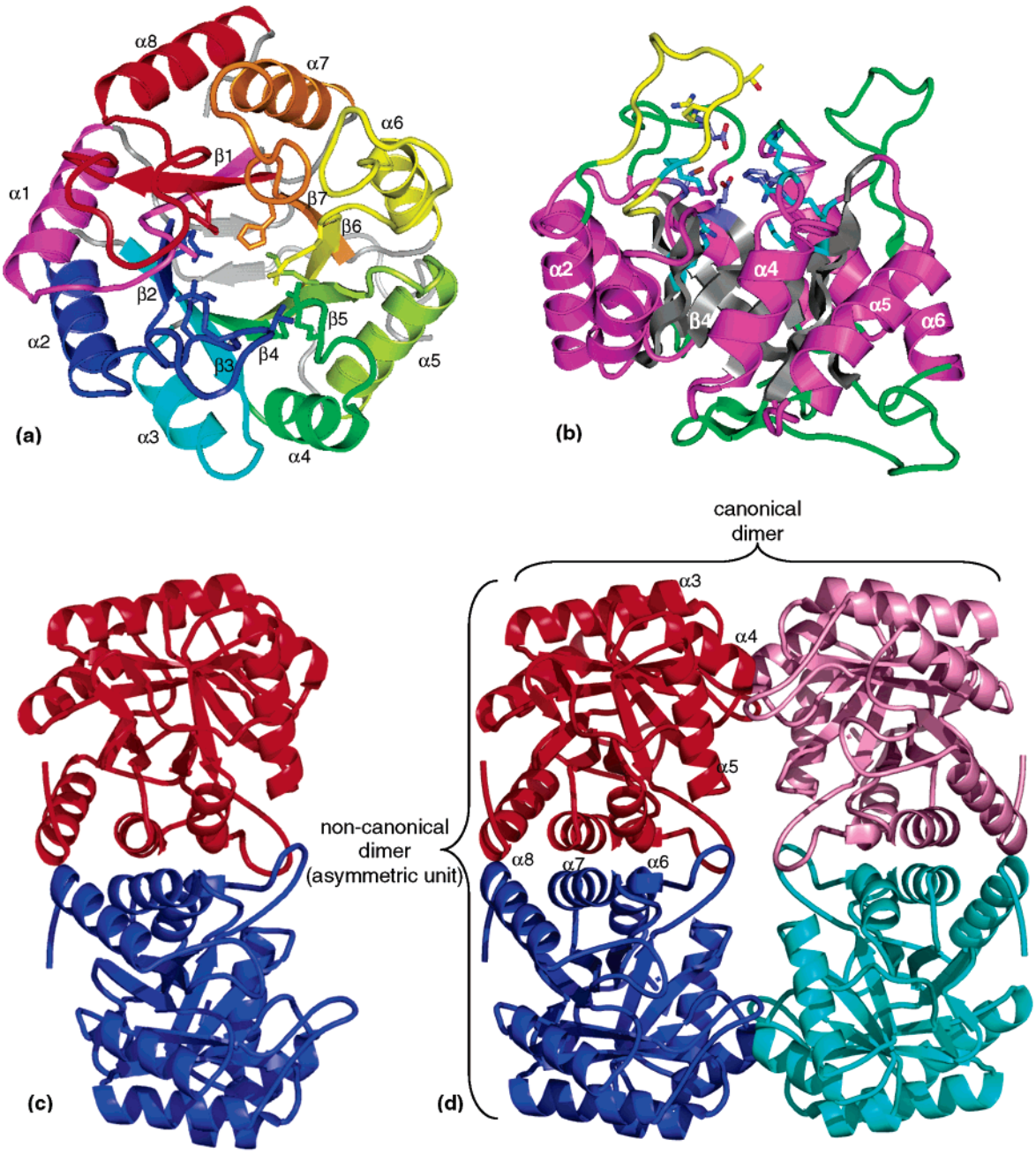



FIGURE 2: *Pf*-DAH7PS monomer, dimer, and tetramer. (a) Rainbow-colored view looking into the $(\beta/\alpha)_8$-barrel, beginning with $\beta$1 and α1 colored violet and ending with $\beta$8 and α8 colored red. Side chains involved in binding PEP and E4P are shown. The α*n*−$\beta$*n*+1 loops and the N-terminal two-stranded $\beta$-hairpin are colored gray. (b) Side-on view with PEP- and E4P-binding loops at the top. The core $(\beta/\alpha)_8$-barrel common to all DAH7PS and KDO8PS enzymes characterized to date is colored gray ($\beta$-strands) and magenta (α-helices). PEP-binding ligands are colored aqua, metal-binding ligands slate, and both the phosphate-binding ligands for the phosphorylated monosaccharide and the $\beta$2−α2 loop yellow. Noncore regions not shared by all type I DAH7PS and KDO8PS enzymes are colored green. (c) Dimer in the asymmetric unit (noncanonical dimer). (d) *Pf*-DAH7PS tetramer.

Metal is found in the active site of subunit B in low occupancy, whereas only traces of metal ions are observed at the metal-binding site of subunit A. The PEP-binding site is partly occupied by PEP, present at relatively low concentrations during crystallization. This leads to some disorder of the side chains that contact PEP. Consistent with crystallization conditions, no phosphorylated monosaccharide is found in the E4P-binding site, and consequently, K60 exhibits considerable disorder.

The other notable difference between the two subunits within the asymmetric unit is the conformation observed for P61. In subunit A, the carbonyl of this residue faces the predicted binding site for E4P in a conformation similar to those described in other DAH7PS structures (*9*, *10*, *13*). In subunit B, however, the carbonyl faces outward, forming a hydrogen bond with a tyrosine from a neighboring subunit in the crystallographic tetramer.

*Quaternary Structure. Pf*-DAH7PS is a dimer in solution as observed by native PAGE and gel filtration chromatography (*29*), yet in the crystal structure, an apparently tight tetrameric association is observed. The asymmetric unit comprises a dimer (termed the noncanonical dimer), formed by contact primarily between helices 6−8 and $\beta$5−α5, $\beta$6−α6, and $\beta$7−α7 loops (Figure 2c). This association buries

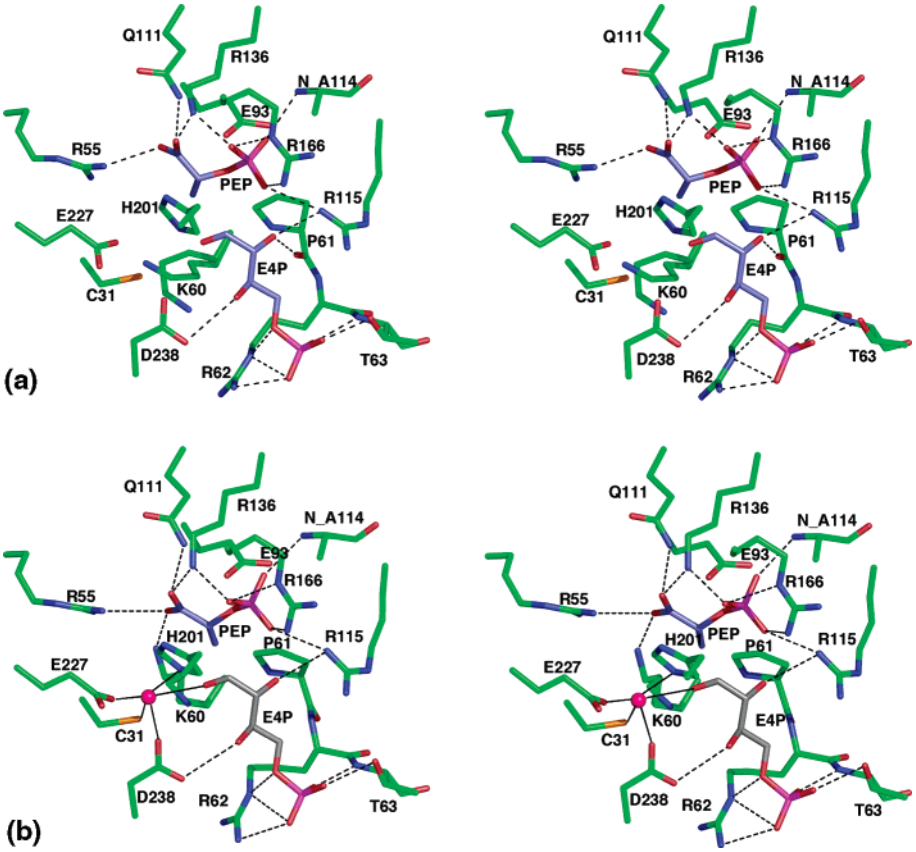



FIGURE 3: Stereodiagram of metal-, PEP-, and E4P-binding ligands in the *Pf*-DAH7PS active site. E4P has been modeled into the structure on the basis of E4P binding to *Tm*-DAH7PS (*13*) and glycerol 3-phosphate binding to *Sc*-DAH7PS (*45*). In other I$\beta$ structures, the amide moiety of the glutamine equivalent to Q111 is rotated 180°. (a) P61 pointing toward modeled E4P as found in subunit A of *Pf*-DAH7PS. (b) P61 pointing away from modeled E4P as found in subunit B of *Pf*-DAH7PS. Metal is present in subunit B, but not in subunit A.

2326 Å$^2$ or 12.2% of the surface area relative to isolated subunits. Four salt bridges, six semisalt bridges ($X^-\cdots H-Y$ or $X-H^+\cdots :Y$) and a single hydrogen bond as well as 11 bridging water molecules comprise the electrostatic component of this interface. A crystallographic 2-fold symmetry operation creates the other interface of the tetramer and involves contacts among $\alpha$4 and $\alpha$5, and $\beta$2−$\alpha$2, $\beta$3−$\alpha$3, $\beta$4−$\alpha$4, $\beta$5−$\alpha$5, and $\beta$6−$\alpha$6 loops (Figure 2d). Relative to isolated subunits, this interface buries 2070 Å$^2$ or 10.8% of the surface. This latter association, which places the active sites close together, is termed the canonical dimer. The canonical dimer has been observed in all DAH7PS and KDO8PS structures characterized to date. For *Pf*-DAH7PS, four salt bridges, 10 semisalt bridges, and four hydrogen bonds, as well as 17 water molecules, render the canonical dimer interface somewhat more electrostatic than the first dimer interface. This second interface, therefore, will be less stable under moderate salt conditions.

*Active Site Organization.* The active site of *Pf*-DAH7PS is found at the C-terminal ends of the $\beta$-strands of the $(\beta/\alpha)_8$-barrel and is formed by residues from both the $\beta n-\alpha n$ loops and $\beta$-strands. Specifically, the first loop, $\beta$1−$\alpha$1, provides metal ligand C31. Strand $\beta$2 and the $\beta$2−$\alpha$2 loop provide PEP- and putative E4P-binding residues R55 and K60, and R62 and T63, respectively. The $\beta$4−$\alpha$4 loop binds PEP through the side chain of R115 and main chain of N_A114. Strand $\beta$5 provides PEP-binding ligand K136. The $\beta$6−$\alpha$6 loop provides PEP-binding ligand R166. The $\beta$7−$\alpha$7 loop provides metal-binding ligand H201. Strand $\beta$8 and the $\beta$8−$\alpha$8 loop provide metal-binding ligands E227 and D238. The metal-, PEP-, and predicted E4P-binding residues are shown in Figure 2; Figure 3 shows the active site in greater detail.

PEP is bound by five arginine and lysine residues that interact with both the phosphate and carboxylate groups (Figure 3). The phosphate group is coordinated by the main chain N atom of A114 and NH1 of R115, NZ of K136, NE and NH2 of R166, and three water molecules. The PEP carboxylate group interacts with NH2 of R55, NZ of K136, NE2 of Q111, and a water molecule. The olefin moiety is in the proximity of a water molecule, which in turn hydrogen bonds to OE2 of E93 in strand $\beta$3.

## DISCUSSION

The wealth of structural data, as well as a growing body of functional data on DAH7PS and KDO8PS proteins, allows new insights into features that are important for ligand binding and reaction mechanism. Crystal structures are now known for two subfamily I$\alpha$ DAH7PS enzymes from *E. coli* and *S. cerevisiae* (*9*, *10*, *42*−*45*), for two subfamily I$\beta_D$ DAH7PS enzymes from *T. maritima* (*13*) and *P. furiosus* (this work), and two subfamily I$\beta_K$ KDO8PS enzymes from *E. coli* and *A. aeolicus* (*11*, *12*, *46*−*49*). All six of these enzymes are homotetramers with $(\beta/\alpha)_8$-barrel subunits containing four apparently independent active sites. However, the considerable differences in quaternary structure and substrate preference provide important clues about the ancestry of this family of proteins.

*Subunit Structure.* Figure 4, a structure-based alignment, highlights the core structural elements of the $(\beta/\alpha)_8$-barrel

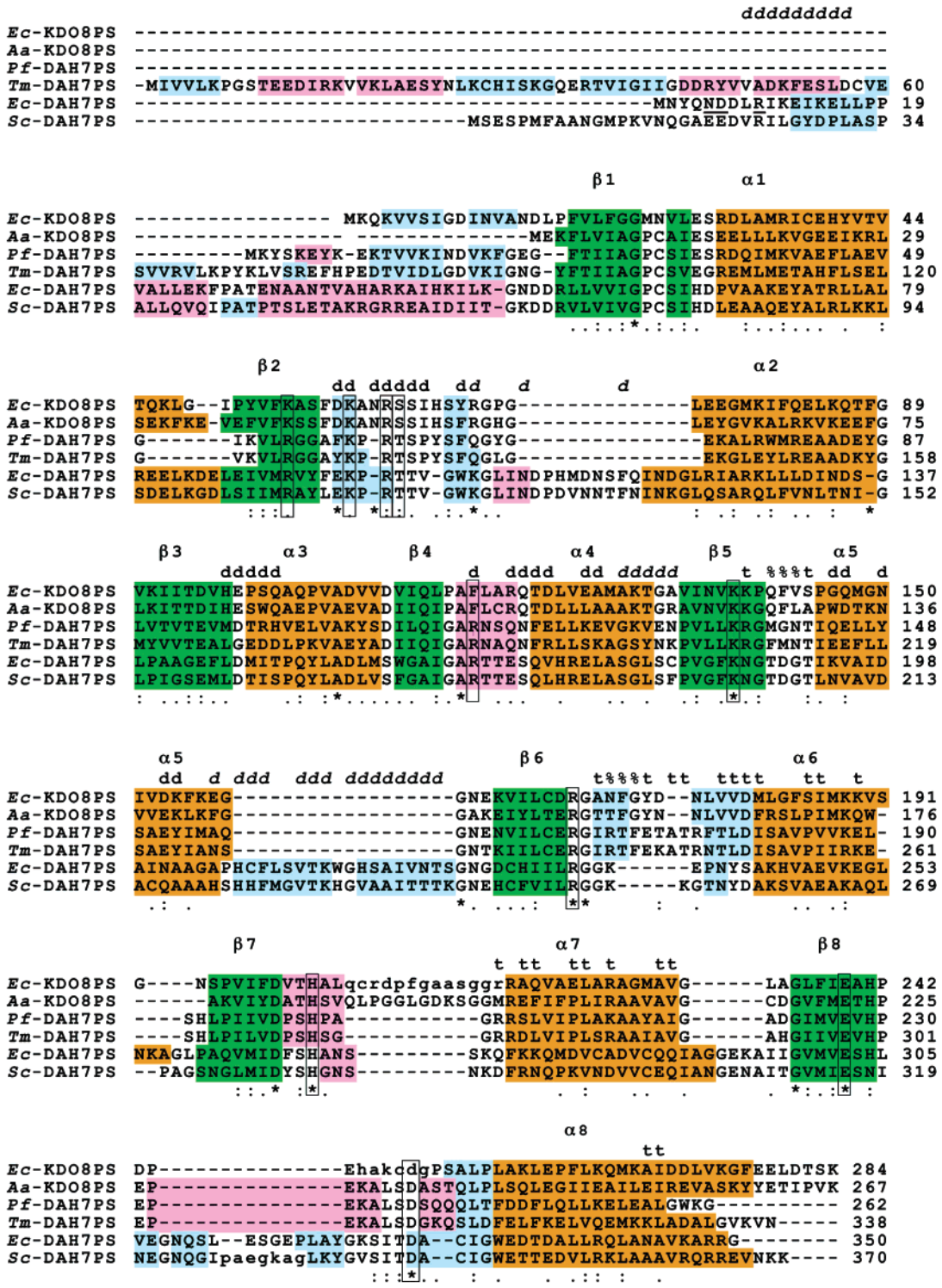



FIGURE 4: Structural alignment of DAH7PS and KDO8PS enzymes. Core $\beta$-strands are highlighted in green and core α-helices in orange. Noncore α-helical regions are highlighted in pink and noncore $\beta$-strands in light blue. Residues involved in the canonical dimer interface are denoted d; those involved in the noncanonical dimer interface are denoted t, and those involved in both interfaces are denoted %. Residues involved in the formation of the noncanonical dimer in Iα DAH7PS enzymes (*Ec*- and *Sc*-DAH7PS) are denoted *d* (in italics). Residues involved in contact with the feedback regulator phenylalanine for the Phe-sensitive isoform of *Ec*-DAH7PS are underlined. Boxes are placed around residues involved in metal, PEP, or phosphorylated monosaccharide binding. Standard ClustalW symbols are used to denote identity (*), a high degree of similarity (:), and a moderate degree of similarity (.).

common to all DAH7PS and KDO8PS enzymes that have been structurally characterized to date (see also Figure 2). The minimalist structure is *Aa*-KDO8PS. Relative to this, the *Ec*-DAH7PS and *Sc*-DAH7PS structures feature 50- and 65-residue N-terminal extensions comprising a $\beta$-strand and a helix−90° turn−helix motif, respectively. Between α5 and $\beta$6, for these two Iα DAH7PS enzymes, an 18-residue extension provides two additional $\beta$-strands. Together with strand $\beta$0, these complete a three-stranded sheet involved in binding of the feedback inhibitor phenylalanine (*Ec*-DAH7PS) or tyrosine (*Sc*-DAH7PS). *Pf*-DAH7PS, *Tm*-DAH7PS, and *Ec*-KDO8PS feature instead a two-stranded

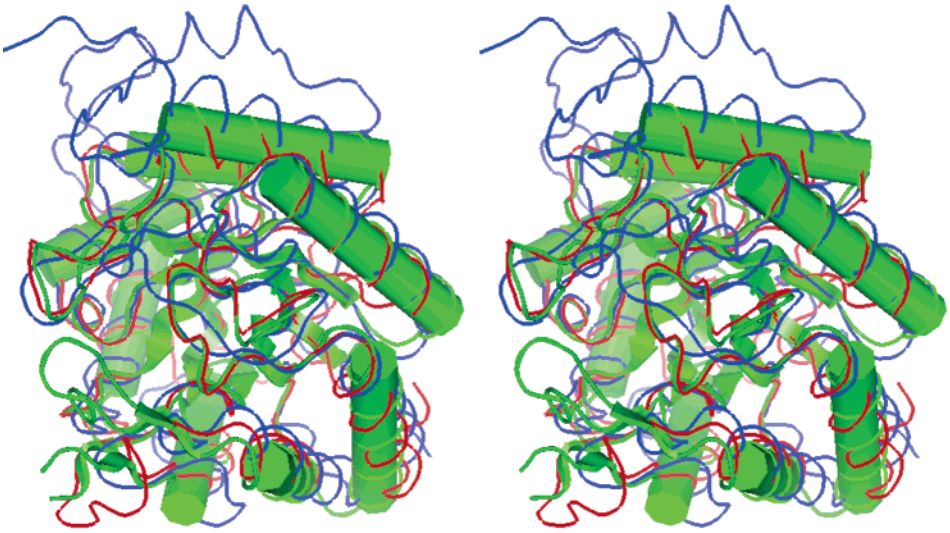


FIGURE 5: Stereodiagram of the superposition of subfamily Iα *Ec*-DAH7PS (blue), subfamily I$\beta_D$ *Pf*-DAH7PS (red), and subfamily I$\beta_K$ *Aa*-KDO8PS (green) structures. To aid visualization, *Aa*-KDO8PS is also shown in cartoon form. The greater structural similarity of subfamily I$\beta_D$ and I$\beta_K$ compared to subfamily Iα is apparent.

$\beta$-hairpin that covers the base of the $(\beta/\alpha)_8$-barrel that is opposite the active site. Spatially, this sheet occupies a similar volume to helix α0B in the Iα DAH7PS enzymes. Relative to *Pf*-DAH7PS, *Tm*-DAH7PS has an additional 78-residue N-terminal ferrodoxin-like domain, which is implicated in feedback inhibition (*13*). *Pf*-DAH7PS and *Tm*-DAH7PS share an abbreviated first helix, α1, compared to both the Iα and I$\beta_K$ enzymes. The $\beta$6−α6 loop forms part of the canonical dimer interface and also forms part of the noncanonical dimer and the tetramer interfaces for subfamily I$\beta$ enzymes. For *Pf*- and *Tm*-DAH7PS, this loop is extended by two residues relative to that of subfamily I$\beta_K$ enzymes and by five or six residues compared to that of subfamily Iα enzymes. Compared to Iα DAH7PS enzymes, *Pf*- and *Tm*-DAH7PS share with subfamily I$\beta_K$ enzymes (i) a shorter $\beta$2−α2 loop (by 14 residues), (ii) a shorter $\beta$8−α8 loop (by ∼15 residues), and (iii) a shorter α7−$\beta$8 loop (by six residues). Relative to subfamily I$\beta_K$ enzymes, Iα and I$\beta_D$ enzymes have a shorter $\beta$7−α7 loop (by nine residues). For KDO8PS, this loop contributes to the dimer interface and provides two additional ligands to the phosphorylated monosaccharide-binding site. To a reasonable approximation, therefore, the *Pf*- and *Tm*-DAH7PS structures resemble more closely the KDO8PS type I$\beta_K$ enzymes. The two regions of difference are (i) the $\beta$7−α7 loop, just mentioned, and (ii) the phosphate-binding motif of the $\beta$2−α2 loop, which contains for KDO8PS the KxxxxKANRS(S/T) motif instead of the RxxxxKPRT(S/T) motif common to all DAH7PS enzymes. Motifs associated with allosteric inhibition of *Sc*- and *Ec*-DAH7PS and *Tm*-DAH7PS are absent in *Pf*-DAH7PS, consistent with the observed absence of feedback inhibition (*29*).

The superposition of subunits of representative Iα (*Ec*-DAH7PS), I$\beta_K$ (*Aa*-KDO8PS), and I$\beta_D$ (*Pf*-DAH7PS) enzymes (Figure 5) shows clearly the greater structural similarity between the subfamily I$\beta$ enzymes than between either I$\beta_D$ or I$\beta_K$ and the Iα subfamily. A threshold of 1.0 Å was set for inclusion in the superposition, to mitigate the effects of loops and to focus on the core structure. Specifically, 228 residues of *Tm*-DAH7PS map onto *Pf*-DAH7PS (62% identical); 131 and 100 residues of *Aa*-KDO8PS (29% identical relative to *Pf*-DAH7PS) and *Ec*-KDO8PS (26% identical), respectively, map similarly, but only 69 and 84 residues of *Ec*-DAH7PS (21%) and *Sc*-DAH7PS (20%), respectively, map. By way of comparison with again a 1.0 Å threshold, for the two Iα structures, 307 residues of *Sc*-DAH7PS map onto *Ec*-DAH7PS (55% identical); for the two I$\beta_K$ structures, 165 residues of *Aa*-KDO8PS map onto *Ec*-KDO8PS (45% identical).

An altered functional role can be ascribed to most of the differences seen in loop regions among the type I DAH7PS and KDO8PS enzymes. (i) Feedback inhibition involves the different N-terminal extensions of *Ec*- and *Sc*-DAH7PS and of *Tm*-DAH7PS, and the extended α5−$\beta$6 loop for *Ec*- and *Sc*-DAH7PS. (ii) Tetrameric association involves the N-terminal extension for *Ec*- and *Sc*-DAH7PS, the extended $\beta$6−α6 loop for subfamily I$\beta$ enzymes, and additionally for *Ec*- and *Aa*-KDO8PS an extended $\beta$7−α7 loop. (iii) Substrate specificity is controlled by the $\beta$2−α2 loop and additionally for *Ec*- and *Aa*-KDO8PS an extended $\beta$7−α7 loop. At this stage, it is not clear what functional role is played by the extended α7−$\beta$8 and $\beta$8−α8 loops in subfamily Iα compared to the subfamily I$\beta$ enzymes.

*Quaternary Structure.* The canonical dimer, which places the E4P-binding $\beta$2−α2 loop of one subunit in contact with the other subunit, is the primary quaternary association observed for the subfamily Iα *Ec*- and *Sc*-DAH7PS. This association is found also in the crystal structures of subfamily I$\beta$ *Aa*- and *Ec*-KDO8PS and *Pf*- and *Tm*-DAH7PS. Significantly different tetrameric associations are apparently characteristic for the two subfamilies. A tetrahedrally twisted tetrameric association is observed crystallographically for subfamily Iα enzymes, and a nearly flat tetrameric association is observed for the subfamily I$\beta$ enzymes.

For all six structures, the canonical interface is made up of contacts among similar structural elements, helices α3−α5 and $\beta$2−α2, $\beta$3−α3, $\beta$4−α4, $\beta$5−α5, and $\beta$6−α6 loops (but not for *Ec*- and *Sc*-DAH7PS, where the $\beta$6−α6 loop is shortened). Despite this, even for the more closely related *Pf*- and *Tm*-DAH7PS, no salt bridges are conserved. For subfamily I$\beta$, the surface areas buried upon formation of the canonical dimer range from ∼2100 to 2500 Å$^2$ (or from 9

to 12% of the total surface area relative to the isolated monomers). For *Ec*- and *Sc*-DAH7PS structures, the canonical dimer interface includes additionally an N-terminal extension, in particular the N-terminal coil and strand $\beta$0, that results in a very large buried surface area of more than 3650 Å$^2$ (or >15% of the total surface area relative to the isolated monomers).

Interaction between subunits to form the canonical dimer interface affects the conformation of a number of key residues involved in ligand binding. In the vicinity of the active site, for *Ec*-KDO8PS (and *Aa*-KDO8PS), R120 (or R106 for *Aa*-KDO8PS) from the $\beta$4−α4 helical turn projects across the canonical dimer interface to hydrogen bond with the main chain carbonyl of N62 (or N48) so that the side chain of this asparagine residue extends into the A5P-binding site (*11*, *12*, *46*). For *Ec*-DAH7PS (and *Sc*-DAH7PS), Q170 (Q185) from the neighboring subunit hydrogen bonds to the main chain carbonyl oxygen of K97 (K112) and the main chain amide nitrogen of R99 (R114) to lock the intervening proline P98 (P113) into a conformation that projects its carbonyl oxygen atom into the E4P-binding pocket (Figure 6c) (*42*, *45*). In addition, R173 (R188) from the adjacent subunit hydrogen bonds to the carbonyl oxygen of phosphate-binding residue T100 (T115). These interactions are absent in the *Pf*- and *Tm*-DAH7PS structures, the glutamine being replaced with phenylalanine and the arginine with leucine. In these I$\beta_D$ DAH7PS enzymes, Y152 (Y223 for *Tm*-DAH7PS) from helix α5 of the adjacent subunit hydrogen bonds to O_R62 (O_R133 for *Tm*-DAH7PS) (*13*). In subunit B of *Pf*-DAH7PS, Y152_OH also hydrogen bonds to O_P61 (Figure 6b). On the other hand, in subunit A of *Pf*-DAH7PS (and in both subunits of *Tm*-DAH7PS), O_P61 (O_P132 for *Tm*-DAH7PS) points away from the tyrosine and a water molecule is found in its place (Figure 6a). This places the proline in a conformation similar to that found for the equivalent proline of both *Ec*- and *Sc*-DAH7PS (*9*, *10*). In this conformation, the carbonyl of this proline is predicted to provide hydrogen bonding to the C2 hydroxyl group of substrate E4P (*13*). Thus, on the basis of crystallographic evidence, for subfamily Iα DAH7PS and subfamily I$\beta_K$ KDO8PS enzymes, interactions between subunits of the canonical dimer appear to be carefully choreographed to place a proline carbonyl oxygen and an asparagine side chain, respectively, in a position to interact with the hydroxyl groups of E4P or A5P (*13*, *50*).

Whereas for subfamily Iα and especially for subfamily I$\beta_K$ enzymes the canonical dimer interface is tightly packed with few cavities, for subfamily I$\beta_D$ enzymes there is a large cavity with a volume in excess of 200 Å$^3$ and a surface area of 190 Å$^2$ near the $\beta$2−α2 loop. In addition, *Pf*-DAH7PS has a cavity with a volume of 56 Å$^3$ and a surface area of 77 Å$^2$ at the intersection of the four subunits of the tetramer (see the table in the Supporting Information for details for each enzyme). The implications for substrate specificity of the tightly packed interface of Iα DAH7PS and I$\beta_K$ KDO8PS compared to the loosely packed interface of I$\beta_D$ DAH7PS enzymes are discussed in a later section.

For subfamily Iα enzymes, the association of a pair of canonical dimers into the tetrahedrally twisted tetramer, through mutual contacts of the $\beta$0−α0A loop and of helices α0A and α2, buries relatively small surface areas of 1300−2000 Å$^2$ (∼5−8%) for various pairwise combinations.

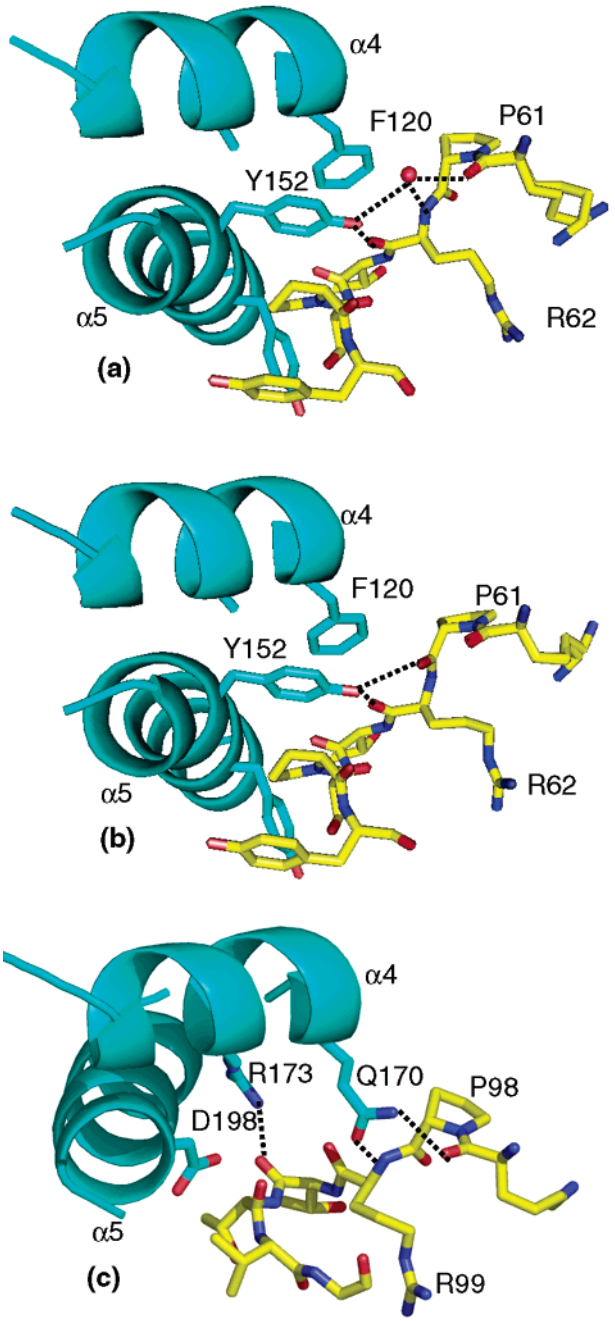



FIGURE 6: Intersubunit contacts in the vicinity of the E4P-binding $\beta$2−α2 loop. A common orientation of helix α4 is used, for which the register of residues among DAH7PS and KDO8PS structures is well conserved compared to helix α5 and the $\beta$2−α2 loop (see Figure 5). (a) *Pf*-DAH7PS subunit A in which Y152 interacts with O_R62 and a water molecule. The *Tm*-DAH7PS canonical dimer interface results in very similar interactions. (b) *Pf*-DAH7PS subunit B in which Y152 interacts with O_R62 and O_P61. This points the carbonyl oxygen atom of P61 away from the predicted E4P binding site. (c) *Ec*-DAH7PS (PDB entry 1gg1) in which Q170 and R173 are found in place of F and Y, respectively, in *Pf*- and *Tm*-DAH7PS structures. The conformation of P98 is locked by tight K97_O···NE2_Q170 and R99_N···OE1_Q170 hydrogen bonds. R99_O hydrogen bonds to R173_NH1 via a water molecule (not shown).

Notwithstanding a seemingly weak dimer-of-dimers tetrameric association, *Ec*-DAH7PS is observed to occur in solution as the tetramer. Although relatively small, the interface between pairs of canonical dimers is tightly packed; no cavities are observed, thereby enhancing the stability of

the tetramer.

In contrast to the Iα enzymes, an apparently tight dimer-of-dimers tetrameric association is observed crystallographically for all subfamily I$\beta$ DAH7PS and KDO8PS proteins structurally characterized to date. This interface is formed by contacts between helices α6−α8 and the $\beta$5−α5, $\beta$6−α6, and $\beta$7−α7 loops. *Ec*- and *Aa*-KDO8PS and *Tm*-DAH7PS are reported to be tetrameric in solution (*11*, *12*, *30*). However, *Pf*-DAH7PS is unequivocally dimeric in solution (*29*), and this dimer is observed to be unusually stable to moderate denaturing conditions. Although surface areas buried in the canonical and noncanonical dimer interfaces are somewhat similar to those of other I$\beta$ tetramers, there are significant differences in the nature of these interfaces. These differences give insight into the apparently anomalous solution-state behavior of *Pf*-DAH7PS. The subfamily I$\beta_K$ enzymes feature tightly packed tetramers with only small cavities, which have surface areas and volumes of less than 110 Å$^2$ and 80 Å$^3$ at subunit interfaces, respectively, consistent with the observed stability of the tetramer in solution (*3*, *32*). Both I$\beta_D$ and I$\beta_K$ enzymes share a tightly packed noncanonical dimer interface. In contrast to enzymes of the I$\beta_K$ and Iα subfamilies, the canonical dimer interface for I$\beta_D$ subfamily enzymes features large cavities.

Focusing on the more closely related subfamily I$\beta_D$ enzymes *Pf*-DAH7PS and *Tm*-DAH7PS, we find both feature a water-filled cavity at the intersection of the four subunits. In addition, the area buried in the noncanonical dimer interface for *Pf*-DAH7PS is greater by ∼300 Å$^2$ than that buried in the canonical dimer interface. In contrast, essentially equal areas are buried in both interfaces for *Tm*-DAH7PS. Moreover, both the total number of contacts and the number of hydrophobic contacts in the canonical dimer interface are substantially greater for *Tm*-DAH7PS (167 and 103, respectively) than for *Pf*-DAH7PS (142 and 86, respectively, for a threshold of 3.8 Å). On the other hand, in the noncanonical dimer interface, these numbers are 100 and 61, and 115 and 71, respectively. In the canonical dimer interface for *Pf*-DAH7PS (*Tm*-DAH7PS in parentheses), there are 4 (6) salt bridges, 10 (4) semisalt bridges, and 4 (2) hydrogen bonds, as well as 17 (20) bridging water molecules. In the noncanonical dimer interface, the numbers are 4 (2), 6 (6), and 1 (1), respectively, with 11 (12) water molecules. Thus, the extent and nature of the subunit interfaces observed crystallographically for *Pf*-DAH7PS are consistent not only with destabilization of this tetramer relative to the *Tm*-DAH7PS tetramer (and to subfamily I$\beta_K$ proteins) but also with the observed solution-state dimer being the noncanonical dimer. The observation of two alternative conformations for P61 in the two subunits of the asymmetric unit for *Pf*-DAH7PS is also consistent with the canonical dimer interface being established on crystallization. The functional unit for *Pf*-DAH7PS is neither the tetramer nor the canonical dimer observed crystallographically in all structures to date, but the noncanonical dimer in which the active sites of the subunits are remote from each other. That a crystallographically observed quaternary association need not correspond to the solution-state quaternary structure, even for substantial buried areas, has been discussed at length and characterized in several systems (*51*−*53*).

*Evolution of Quaternary Structure and Allostery.* Except for residues involved in substrate binding, there is very little conservation of the canonical dimer interface, even between the subfamily I$\beta_K$ and I$\beta_D$ enzymes. On the other hand, the noncanonical dimer interface, although well removed from the active site, is highly conserved, including two salt bridges and several hydrophobic patches. Only 16 of the 27 residues in contact in the canonical dimer interface for *Pf*-DAH7PS are conserved in *Tm*-DAH7PS. Another two residues are very conservatively substituted. Of these 16, 11 are absolutely conserved and involved in substrate binding across all structurally characterized DAH7PS enzymes. At the noncanonical dimer interface, a much higher level of conservation is seen, with 20 of 27 residues conserved and another four very conservatively substituted. In other words, for subfamily I$\beta_D$ enzymes, there appear to be stronger evolutionary constraints to change at the noncanonical dimer interface compared to the canonical dimer interface. Thus, the ancestral DAH7PS would appear to have been the noncanonical dimer, formed by contact of helices and loops remote from the active site and incapable of regulation by allosteric ligand binding. Association of this dimeric structure into tetramers (represented by the contemporary proteins *Aa*-KDO8PS and *Ec*-KDO8PS) was possibly followed by acquisition of N-terminal domains associated with allosteric inhibition (as seen in *Tm*-DAH7PS). Analysis of the N-terminal sequences of *Sc*- and *Ec*-DAH7PS indicates that these may be truncated, and differently folded, versions of the ferrodoxin-like domain attached to *Tm*-DAH7PS. This alternative mode of allosteric inhibition in these more evolutionarily remote Iα enzymes (*Sc*-DAH7PS and *Ec*-DAH7PS) is associated also with the addition of $\beta$-strands between helices α5 and $\beta$6 and the loss of the *Pf*-DAH7PS-like noncanonical dimer.

*Active Site Structure.* All residues that interact with PEP in the active sites of other DAH7PS enzymes are conserved in *Pf*-DAH7PS, with one exception. For *Pf*- and *Tm*-DAH7PS (and also for *Ec*- and *Aa*-KDO8PS), the NE2 atom of Q111 and Q186, respectively, from the $\beta$4−α4 helical loop hydrogen bonds to the PEP carboxylate group. In contrast, for *Ec*- and *Sc*-DAH7PS, a water molecule is bound, which in turn hydrogen bonds to a tyrosine (Y94 and Y109, respectively) from the $\beta$2 strand. While the majority of PEP-coordinating contacts in *Pf*-DAH7PS and *Aa*-KDO8PS are also similar, several differences are observed. The arginine on the $\beta$4−α4 helical loop (R115 for *Pf*-DAH7PS), which binds the phosphate group of PEP, is absent in KDO8PS proteins. In addition, the interaction of the PEP carboxylate group with NH2 of R55 on the $\beta$2−α2 loop is replaced with a salt bridge with NZ of a structurally equivalent lysine, and a water molecule is replaced with OG of serine on the $\beta$7−α7 loop in *Aa*-KDO8PS.

The metal-binding site is also similar to that found in other metal-dependent enzymes of this family (*Ec*-DAH7PS, *Sc*-DAH7PS, *Tm*-DAH7PS, and *Aa*-KDO8PS). In the metal-independent enzyme, *Ec*-KDO8PS, the metal-binding site is filled with an asparagine side chain in place of C31 and the aspartate ligand is flipped away from the histidine and glutamate side chains, possibly due in part to a proline two residues downstream. This proline is absent in metal-binding DAH7PS and KDO8PS.

There are several places where motifs associated with E4P-processing DAH7PS enzymes, rather than A5P-processing KDO8PS enzymes, are found. The most obvious motif is the RxxxxKPRT motif on the $\beta$2−α2 loop of DAH7PS,

which for A5P-processing KDO8PS enzymes occurs as KxxxxKANRS. A second site, glutamate (E93 for *Pf*-DAH7PS) on strand $\beta$3 (D in KDO8PS), is coupled with (I/L)GAR on the $\beta$4−α4 loop [(I/L)PAF in KDO8PS], where the glycine and proline control positioning of the glutamate and aspartate side chains, respectively.

*Implications of Structure for Substrate Specificity.* Analysis of the substrate specificity of *Pf*-DAH7PS shows that this enzyme does not catalyze the aldol-like reaction between PEP and phosphorylated monosaccharide substrates one carbon atom shorter or two carbon atoms longer than the natural substrate, E4P. On the other hand, all five-carbon phosphorylated monosaccharides that were tested were substrates. Enzymic activity with five-carbon phosphorylated sugars has also been reported for subfamily Iα DAH7PS-(phe) from *E. coli*, but A5P, 2dR5P, and R5P are considerably poorer substrates for this enzyme than for *Pf*-DAH7PS. For *Ec*-DAH7PS, $k_{cat}$ values range between 0.5 and 7% when compared to that for the use of E4P (*28*), whereas for *Pf*-DAH7PS, the range is 76−177%. *Ec*-DAH7PS also exhibited no activity with G3P or G6P (*28*). In contrast, the only other characterized DAH7PS from subfamily I$\beta$, *Tm*-DAH7PS, is reported to show no enzymic activity with either A5P or R5P (*30*). While the apparent $K_M$ values recorded in this study for the five-carbon substrates are relatively large compared to the $K_M^{E4P}$ value, it is noteworthy that the $k_{cat}$ values are comparable to that of E4P. This suggests that these molecules bind to the enzyme active site forming a complex that has catalytic competency equivalent to that of the enzyme−PEP−E4P complex. As A5P, 2dR5P, and R5P all exist predominantly in cyclic forms in solution, and all evidence to date suggests that the free aldehyde form is utilized in the enzyme-catalyzed reaction, it is likely that the true $K_M$ values for these five-carbon substrates are considerably smaller than the apparent $K_M$ values reported in this study.

Remarkably, the four-carbon E4P analogue 2dE4P is a better substrate for *Pf*-DAH7PS than E4P itself. The efficiency with which this compound is utilized indicates that the C2 hydroxyl group of E4P does not play a significant role in substrate binding and catalysis in this enzyme. On the other hand, 2dE4P is a relatively poor substrate of *Ec*-DAH7PS ($K_M^{2dE4P} = 407\ \mu$M and $k_{cat} = 18\ s^{-1}$, compared to $K_M^{E4P} = 21\ \mu$M and $k_{cat} = 71\ s^{-1}$, unpublished results).

Analysis of the structure of *Pf*-DAH7PS provides clues about the differences in substrate specificity between *Pf*-DAH7PS and the *Tm*- and *Ec*-DAH7PS enzymes. On the basis of previous structural work, the E4P binding site for all these enzymes is provided chiefly by residues on the $\beta$2−α2 loop (*13*, *45*). Intersubunit interactions that form the canonical dimer place significant constraints on the conformation of this loop. Whereas *Tm*- and *Ec*-DAH7PS are tetrameric in solution, the solution structure of *Pf*-DAH7PS is dimeric and as discussed above most likely the noncanonical dimer. Consequently, the intersubunit interactions that orchestrate the placement of side chains to form the phosphate-binding site and to provide hydrogen-bonding contacts to the hydroxyl groups for E4P in other DAH7PS enzymes are absent in *Pf*-DAH7PS in solution. The extra flexibility in this part of the active site may explain the relatively high $k_{cat}$ values for the use of five-carbon substrates observed for *Pf*-DAH7PS compared to *Ec*-DAH7PS and *Tm*-DAH7PS. As noted above, P61 in subunit B of *Pf*-DAH7PS adopts a conformation not previously observed for DAH7PS, where the carbonyl points away from the predicted E4P binding site. P61 is part of the absolutely conserved KPRT motif observed in all type I DAH7PS enzymes. Models of E4P in the active sites of *Ec*-, *Sc*-, and *Tm*-DAH7PS suggest that the main chain carbonyl of the proline in this motif hydrogen bonds with the C2 hydroxyl group of E4P (*13*, *45*). The altered conformation of this proline in *Pf*-DAH7PS, or at least increased flexibility in its positioning, may contribute to 2dE4P being utilized more efficiently than natural substrate E4P. Studies with site-directed mutants are currently underway to establish unequivocally the nature of the dimeric species observed in solution, and the influence of dimer interactions on substrate specificity.

KDO8PS enzymes also exhibit relatively tight substrate specificity. E4P and R5P are not substrates for either *Ec*- or *Aa*-KDO8PS, although 2dR5P is reported to show limited substrate activity with *Ec*-KDO8PS (*33*). The phosphorylated monosaccharide binding sites for these proteins, also formed in large part by residues of the $\beta$2−α2 loop, are also expected to be significantly buttressed by the formation of the canonical dimer.

The insensitivity of *Pf*-DAH7PS to changes at the C2 position of the phosphorylated monosaccharide may be indicative of a relationship between this archaeal enzyme and KDO8PS, which uses A5P as a substrate. The efficiency of 2dE4P as a substrate and the ability of the enzyme to accept five-carbon substrates in which the C2 hydroxyl group is either absent or present in either possible configuration suggest that the C2 hydroxyl group does not have a significant interaction with this enzyme. In contrast, the correct stereochemistry of the C2 hydroxyl group appears to be required in KDO8PS (subfamily I$\beta_K$), with the natural substrate A5P having the opposite configuration at C2 compared to E4P. Subramaniam et al. (*14*) proposed that the ancestral protein for this enzyme family is a DAH7PS. The substrate ambiguity that we have observed for *Pf*-DAH7PS further suggests that in developing KDO8PS activity the ancestral 3-deoxyald-2-ulosonate-phosphate synthase evolved by losing the ability to accept four-carbon monosaccharides and by a narrowing specificity at C2. This broad substrate specificity, the requirement for a divalent metal ion, the lack of allosteric inhibition (*29*), the dimeric solution-state quaternary structure, and the proposal that the initial ancestor was most likely the I$\beta$ type of DAH7PS (*14*, *15*) indicate that DAH7PS from *P. furiosus* may be the contemporary protein that is most similar to the ancestral type I DAH7PS, or 3-deoxyald-2-ulosonate-phosphate synthase.

## ACKNOWLEDGMENT

We thank Amy Pietersma for preparation of 2dE4P.

## SUPPORTING INFORMATION AVAILABLE

A table of calculated surface areas and cavities for *Pf*-DAH7PS, *Tm*-DAH7PS, *Ec*-DAH7PS, *Sc*-DAH7PS, *Aa*-KDO8PS, and *Ec*-KDO8PS and a figure highlighting the different intersubunit contacts and quaternary assemblies for *Pf*-DAH7PS, *Tm*-DAH7PS, *Aa*-KDO8PS, and *Ec*-DAH7PS.

This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES


1. Bentley, R. (1990) The shikimate pathway: A metabolic tree with many branches, *Crit. Rev. Biochem. Mol. Biol. 25*, 307−384.
2. Herrmann, K. M., and Weaver, L. M. (1999) The shikimate pathway, *Annu. Rev. Plant Physiol. Plant Mol. Biol. 50*, 473−503.
3. Ray, P. H. (1980) Purification and characterization of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthetase from *Escherichia coli*, *J. Bacteriol. 141*, 635−644.
4. Raetz, C. R. (1990) Biochemistry of endotoxins, *Annu. Rev. Biochem. 59*, 129−170.
5. DeLeo, A. B., Dayan, J., and Sprinson, D. B. (1973) Purification and kinetics of tyrosine-sensitive 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate synthetase from *Salmonella*, *J. Biol. Chem. 248*, 2344−2353.
6. DeLeo, A. B., and Sprinson, D. B. (1968) Mechanism of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate (DAHP) synthetase, *Biochem. Biophys. Res. Commun. 32*, 873−877.
7. Onderka, D. K., and Floss, H. G. (1969) Steric course of the chorismate synthetase reaction and the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthetase reaction, *J. Am. Chem. Soc. 91*, 5894−5896.
8. Kohen, A., Berkovich, R., Belakhov, V., and Baasov, T. (1993) Stereochemistry of the KDO8P synthase. An efficient synthesis of the 3-fluoro analogs of KDO8P, *Bioorg. Med. Chem. Lett. 3*, 1577−1582.
9. Shumilin, I. A., Kretsinger, R. H., and Bauerle, R. H. (1999) Crystal structure of phenylalanine-regulated 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*, *Structure 7*, 865−875.
10. Hartmann, M., Schneider, T. R., Pfeil, A., Heinrich, G., Lipscomb, W. N., and Braus, G. H. (2003) Evolution of feedback-inhibited $\beta/\alpha$ barrel isoenzymes by gene duplication and a single mutation, *Proc. Natl. Acad. Sci. U.S.A. 100*, 862−867.
11. Duewel, H. S., Radaev, S., Wang, J., Woodard, R. W., and Gatti, D. L. (2001) Substrate and metal complexes of 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase from *Aquifex aeolicus* at 1.9 Å resolution. Implications for the condensation mechanism, *J. Biol. Chem. 276*, 8393−8402.
12. Radaev, S., Dastidar, P., Patel, M., Woodard, R. W., and Gatti, D. L. (2000) Structure and mechanism of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase, *J. Biol. Chem. 275*, 9476−9484.
13. Shumilin, I. A., Bauerle, R., Wu, J., Woodard, R. W., and Kretsinger, R. H. (2004) Crystal structure of the reaction complex of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Thermotoga maritima* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation, *J. Mol. Biol. 341*, 455−466.
14. Subramaniam, P. S., Xie, G., Xia, T., and Jensen, R. A. (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.
15. Jensen, R. A., Xie, G., Calhoun, D. H., and Bonner, C. A. (2002) 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), *J. Mol. Evol. 54*, 416−423.
16. Birck, M. R., and Woodard, R. W. (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.
17. Walker, G. E., Dunbar, B., Hunter, I. S., Nimmo, H. G., and Coggins, J. R. (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.
18. Paravicini, G., Schmidheini, T., and Braus, G. (1989) Purification and properties of the 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (phenylalanine-inhibitable) of *Saccharomyces cerevisiae*, *Eur. J. Biochem. 186*, 361−366.
19. Schnappauf, G., Hartmann, M., Kunzler, M., and Braus, G. H. (1998) The two 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase isoenzymes from *Saccharomyces cerevisiae* show different kinetic modes of inhibition, *Arch. Microbiol. 169*, 517−524.
20. Hoffmann, P. J., Doy, C. H., and Catcheside, D. E. A. (1972) Separation of three allosterically inhibitable 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthases from extracts of *Neurospora crassa* and the purification of the tyrosine inhibitable isoenzyme, *Biochim. Biophys. Acta 268*, 550−561.
21. Nimmo, G. A., and Coggins, J. R. (1981) The purification and molecular properties of the tryptophan-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Neurospora crassa*, *Biochem. J. 197*, 427−436.
22. Nimmo, G. A., and Coggins, J. R. (1981) Some kinetic properties of the tryptophan-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Neurospora crassa*, *Biochem. J. 199*, 657−665.
23. Gosset, G., Bonner, C. A., and Jensen, R. A. (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.
24. Schoner, R., and Herrmann, K. M. (1976) 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. Purification, properties, and kinetics of the tyrosine-sensitive isoenzyme from *Escherichia coli*, *J. Biol. Chem. 251*, 5440−5447.
25. McCandliss, R. J., Poling, M. D., and Herrmann, K. M. (1978) 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. Purification and molecular characterization of the phenylalanine-sensitive isoenzyme from *Escherichia coli*, *J. Biol. Chem. 253*, 4259−4265.
26. Ray, J. M., and Bauerle, R. (1991) Purification and properties of tryptophan-sensitive 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*, *J. Bacteriol. 173*, 1894−1901.
27. Stephens, C. M., and Bauerle, R. (1991) Analysis of the metal requirement of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*, *J. Biol. Chem. 266*, 20810−20817.
28. Sheflyan, G. Y., Howe, D. L., Wilson, T. L., and Woodard, R. W. (1998) Enzymic synthesis of 3-deoxy-D-*manno*-octulosonate 8-phosphate, 3-deoxy-D-*altro*-octulosonate 8-phosphate, 3,5-dideoxy-D-*gluco*(*manno*)-octulosonate 8-phosphate by 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase, *J. Am. Chem. Soc. 120*, 11027−11032.
29. Schofield, L. R., Patchett, M. L., and Parker, E. J. (2004) Expression, purification, and characterization of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Pyrococcus furiosus*, *Protein Expression Purif. 34*, 17−27.
30. Wu, J., Howe, D. L., and Woodard, R. W. (2003) *Thermotoga maritima* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthase: The ancestral eubacterial DAHP synthase? *J. Biol. Chem. 278*, 27525−27531.
31. Duewel, H. S., and Woodard, R. W. (2000) A metal bridge between two enzyme families. 3-Deoxy-D-*manno*-octulosonate-8-phosphate synthase from *Aquifex aeolicus* requires a divalent metal for activity, *J. Biol. Chem. 275*, 22824−22831.
32. Duewel, H. S., Sheflyan, G. Y., and Woodard, R. W. (1999) Functional and biochemical characterization of a recombinant 3-deoxy-D-*manno*-octulosonic acid 8-phosphate synthase from the hyperthermophilic bacterium *Aquifex aeolicus*, *Biochem. Biophys. Res. Commun. 263*, 346−351.
33. Howe, D. L., Sundaram, A. K., Wu, J., Gatti, D. L., and Woodard, R. W. (2003) Mechanistic insight into 3-deoxy-D-*manno*-octulosonate-8-phosphate synthase and 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase utilizing phosphorylated monosaccharide analogues, *Biochemistry 42*, 4843−4854.
34. Sheflyan, G. Y., Sundaram, A. K., Taylor, W. P., and Woodard, R. W. (2000) Substrate ambiguity of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* revisited, *J. Bacteriol. 182*, 5005−5008.
35. Fiala, G., and Stetter, K. O. (1986) *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100 °C, *Arch. Microbiol. 145*, 56−61.
36. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite, *Acta Crystallogr. B59*, 1131−1137.
37. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server, *Bioinformatics 16*, 404−405.
38. McGuffin, L. J., and Jones, D. T. (2003) Improvement of the GenTHREADER method for genomic fold recognition, *Bioinformatics 19*, 874−881.

39. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) SWISS-MODEL: An automated protein homology-modeling server, *Nucleic Acids Res. 31*, 3381−3385.
40. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins: Struct., Funct., Genet. 11*, 281−296.
41. DeLano, W. L. (2002) *The PyMol molecular graphics system*, DeLano Scientific, San Carlos, CA.
42. Wagner, T., Shumilin, I. A., Bauerle, R., and Kretsinger, R. H. (2000) 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-phosphoenolpyruvate complexes and implications for catalysis, *J. Mol. Biol. 301*, 389−399.
43. Shumilin, I. A., Bauerle, R., and Kretsinger, R. H. (2003) The high-resolution structure of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase reveals a twist in the plane of bound phosphoenolpyruvate, *Biochemistry 42*, 3766−3776.
44. Shumilin, I. A., Zhao, C., Bauerle, R., and Kretsinger, R. H. (2002) Allosteric inhibition of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase alters the coordination of both substrates, *J. Mol. Biol. 320*, 1147−1156.
45. Konig, V., Pfeil, A., Braus, G. H., and Schneider, T. R. (2004) Substrate and metal complexes of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Saccharomyces cerevisiae* provide new insights into the catalytic mechanism, *J. Mol. Biol. 337*, 675−690.
46. Wagner, T., Kretsinger, R. H., Bauerle, R., and Tolbert, W. D. (2000) 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.
47. Asojo, O., Friedman, J., Adir, N., Belakhov, V., Shoham, Y., and 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.
48. Wang, J., Duewel, H. S., Woodard, R. W., and Gatti, D. L. (2001) Structures of *Aquifex aeolicus* KDO8P synthase in complex with R5P and PEP, and with a bisubstrate inhibitor: Role of active site water in catalysis, *Biochemistry 40*, 15676−15683.
49. Wang, J., Duewel, H. S., Stuckey, J. A., Woodard, R. W., and Gatti, D. L. (2002) Function of His185 in *Aquifex aeolicus* 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase, *J. Mol. Biol. 324*, 205−214.
50. Williamson, R. M., Pietersma, A. L., Jameson, G. B., and Parker, E. J. (2005) Stereospecific deuteration of 2-deoxyerythrose 4-phosphate using 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase, *Bioorg. Med. Chem. Lett. 15*, 2339−2342.
51. Bahadur, R. P., Chakrabarti, P., Rodier, F., and Janin, J. (2004) A dissection of specific and non-specific protein−protein interfaces, *J. Mol. Biol. 336*, 943−955.
52. Valdar, W. S. J., and Thornton, J. M. (2001) Conservation helps to identify biologically relevant crystal contacts, *J. Mol. Biol. 313*, 399−416.
53. Ding, H., Hunt, J. F., Mukerji, I., and Oliver, D. (2003) *Bacillus subtilis* SecA ATPase exists as an antiparallel dimer in solution, *Biochemistry 42*, 8729−8738.

BI050577Z