

Research Article

Structural and functional characterisation of the entry point to pyocyanin biosynthesis in *Pseudomonas aeruginosa* defines a new 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase subclass

Oliver W. Sterritt^{1,2}, Eric J.M. Lang^{1,*}, Sarah A. Kessans¹, Timothy M. Ryan³,  Borries Demeler⁴,  Geoffrey B. Jameson^{1,2,5} and  Emily J. Parker^{1,2,6}

¹Biomolecular Interaction Centre and School of Physical and Chemical Sciences, University of Canterbury, Christchurch 8041, New Zealand; ²Maurice Wilkins Centre for Molecular Biodiscovery, University of Auckland, Auckland 1142, New Zealand; ³SAXS/WAXS, Australian Synchrotron/ANSTO, 800 Blackburn Road, Clayton, VIC 3168, Australia; ⁴Department of Chemistry and Biochemistry, University of Lethbridge, Alberta T1K 3M4, Canada; ⁵Institute of Fundamental Sciences and the Riddet Institute, Massey University, Palmerston North 4442, New Zealand; ⁶Ferrier Research Institute, Victoria University of Wellington, Wellington 6140, New Zealand

Correspondence: Emily J. Parker (emily.parker@vuw.ac.nz) or Geoffrey B. Jameson (g.b.jameson@massey.ac.nz)



In *Pseudomonas aeruginosa* (*Pae*), the shikimate pathway end product, chorismate, serves as the last common precursor for the biosynthesis of both primary aromatic metabolites, including phenylalanine, tyrosine and tryptophan, and secondary aromatic metabolites, including phenazine-1-carboxylic acid (PCA) and pyocyanin (PYO). The enzyme 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAH7PS) catalyses the first committed step of the shikimate pathway, en route to chorismate. *P. aeruginosa* expresses multiple, distinct DAH7PSs that are associated with either primary or secondary aromatic compound biosynthesis. Here we report the structure of a type II DAH7PS, encoded by *phzC* as part of the duplicated phenazine biosynthetic cluster, from *P. aeruginosa* (PAO1) revealing for the first time the structure of a type II DAH7PS involved in secondary metabolism. The omission of the structural elements α_{2a} and α_{2b} , relative to other characterised type II DAH7PSs, leads to the formation of an alternative, dimeric, solution-state structure for this type II DAH7PS with an oligomeric interface that has not previously been characterised and that does not facilitate the formation of aromatic amino acid allosteric binding sites. The sequence similarity and, in particular, the common N-terminal extension suggest a common origin for the type II DAH7PSs from *P. aeruginosa*. The results described in the present study support an expanded classification of the type II DAH7PSs as type II_A and type II_B based on sequence characteristics, structure and function of the resultant proteins, and on defined physiological roles within primary or secondary metabolism.

* Present address: School of Chemistry, University of Bristol, Bristol BS8 1TS, U.K.

Received: 22 July 2018
Revised: 17 September 2018
Accepted: 20 September 2018

Accepted Manuscript Online:
21 September 2018
Version of Record published:
17 October 2018

Introduction

Pseudomonas aeruginosa (*Pae*) is an opportunistic human pathogen often associated with the chronic infection of patients suffering from cystic fibrosis [1]. *P. aeruginosa* produces a number of virulence factors [2] that are involved in bacterial quorum sensing [3] and establishing long-term infections, particularly of the lungs. Pyocyanin (PYO) is a secondary metabolite derived from phenazine-1-carboxylic acid (PCA) that has been shown to interfere with a number of cellular processes [4-7] and is an essential virulence factor required for pathogenic infection [8].

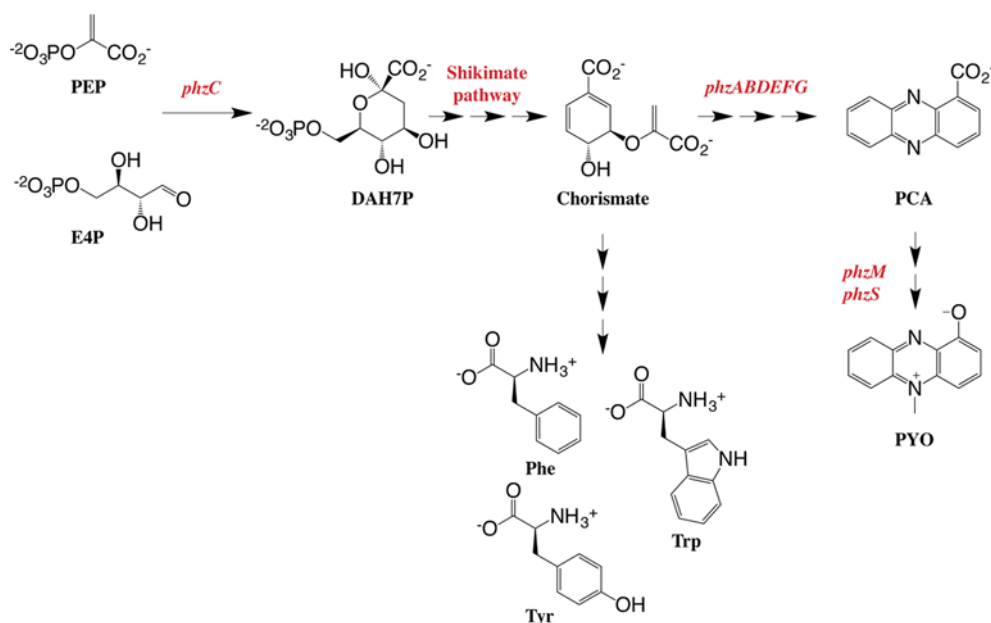


Figure 1. The shikimate pathway is responsible for the formation of aromatic compounds in microorganisms and plants

In case of *P. aeruginosa*, the pathway end product, chorismate, is the last common precursor for the biosynthesis of both primary and secondary aromatic metabolites.

In *P. aeruginosa*, the shikimate pathway end product, chorismate, acts as the last common precursor for both primary aromatic metabolism, for the biosynthesis of the aromatic amino acids phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp), and secondary aromatic metabolism including the biosynthesis of PCA and PYO. Starting from chorismate, the first steps of PYO biosynthesis are carried out by the duplicated seven-gene operon *phzABCDEFGHI* [9], leading to the formation of PCA. Expression of the *phzA-G* operon is under genetic control by the LysR-like transcriptional regulator *mvfR* [10,11] as part of the *P. aeruginosa* quorum-sensing circuitry. PCA is converted into PYO through additional two steps (*phzM* and *phzS*) (Figure 1) [12].

The enzymes of the *phzA-G* biosynthetic operon, along with the enzymes required for the biosynthesis of PYO from PCA, have been well characterised [13-19], with the exception of the enzyme derived from *phzC*. *PhzC* encodes a putative 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAH7P) synthase (DAH7PS), which catalyses the aldol-like condensation reaction between phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) to form DAH7P as the first committed step of the shikimate pathway, en route to chorismate.

DAH7PSs have been classified into three broad groupings based on enzyme sequence: type I α , type I β and type II [20,21]. Although less than 10% sequence identity exists between the type I and II DAH7PS groupings, all characterised examples of DAH7PSs share a common (β/α)₈-barrel fold, a common divalent metal-ion binding site and conservation of almost all of the residues involved with E4P and PEP binding [22-33]. Various structural elements, additional to the core catalytic barrel, are associated with a diverse set of allosteric responses and the formation of alternate quaternary assemblies. The nature and location of these additional structural elements within the core catalytic barrel is characteristic of each group of DAH7PS enzymes.

While the characteristics of many examples of type I DAH7PSs have been reported, characterisation of the type II DAH7PSs has focused primarily on a group of type II enzymes that, relative to the minimalist type I β unadorned catalytic barrels such as *Pyrococcus furiosus* DAH7PS [25], contain both an approximately 75-residue N-terminal extension (typically providing elements β_0 , α_{0a} , α_{0b} and α_{0c}) and an approximately 60-residue extension to loop $\alpha_2\beta_3$ (typically providing elements α_{2a} and α_{2b}). For example, *Mycobacterium tuberculosis* (*Mtu*) expresses a single type II DAH7PS (*Mtu*DAH7PS), which contains these accessory structural elements. The extra-barrel elements in *Mtu*DAH7PS provide three distinct allosteric binding sites, on the single enzyme, that are each selective for either Trp, Tyr or Phe, and together they contribute towards a complex allosteric regulatory mechanism where binary or ternary combinations of aromatic amino acids that include Trp act synergistically to inhibit the enzyme [34-36]. These extensions are also responsible for the formation of the oligomeric interfaces that are present in the homotetrameric assemblies of the characterised type II enzymes. The allosteric functionality of either *Mtu*DAH7PS or the type II DAH7PS

from *Corynebacterium glutamicum* (CglDAH7PS) is extended by the formation of a non-covalent complex with the AroQ₅ subclass of chorismate mutase (*Mtu*CM or *Cgl*CM respectively). The formation of this non-covalent complex results in an activity boost for the CM while allowing the CM to access and utilise the allosteric machinery located on the DAH7PS [32,37,38].

In comparison, *P. aeruginosa* expresses two type I α and two type II DAH7PSs. The type II DAH7PSs are encoded by the ORFs PA1901 (and duplicated as PA4212) and PA2843 (*Pae*DAH7PS^{PA1901} and *Pae*DAH7PS^{PA2843} respectively). The structure and properties of *Pae*DAH7PS^{PA2843} have recently been reported [33] and show that *Pae*DAH7PS^{PA2843} contains an N-terminal extension that is 19 residues shorter in sequence length and has similar inserted α_{2a} and α_{2b} helices, as compared with *Mtu*DAH7PS or *Cgl*DAH7PS. Although the quaternary assemblies of *Mtu*DAH7PS and *Pae*DAH7PS^{PA2843} resemble each other, there are some differences in orientation of the extra-barrel elements within the tetramer, such that *Pae*DAH7PS^{PA2843} is inhibited by only Trp and is unaffected by combinations of Trp with Phe and Tyr due to the abbreviated N-terminal extension.

However, based on sequence alignments, there appears to be a second distinct grouping of type II DAH7PSs [39]. These DAH7PSs appear to contain an N-terminal extension to the core (β/α)₈-barrel, comparable with that observed in *Pae*DAH7PS^{PA2843}, but lack sequence corresponding to the inserted helices α_{2a} and α_{2b} for *Pae*DAH7PS^{PA2843}. This difference in accessory structural elements may therefore have consequences for the formation of the quaternary assembly and hence for allosteric functionality. For example, *Pae*DAH7PS^{PA1901}, found as a part of the *phzA*-G PCA biosynthetic cluster, is 43 amino acids shorter in sequence length relative to *Pae*DAH7PS^{PA2843}, which is not found within the *phzA*-G PCA biosynthetic cluster. A primary structure alignment among type II DAH7PSs (Supplementary Figure S1) indicates that *Pae*DAH7PS^{PA1901} appears to lack much of the sequence that corresponds to the α_{2a} and α_{2b} helices of *Pae*DAH7PS^{PA2843}, *Mtu*DAH7PS or *Cgl*DAH7PS, and hence belongs to this structurally and functionally uncharacterised second group of type II DAH7PSs.

Here we report the structure of *Pae*DAH7PS^{PA1901}, characterising for the first time a short-form type II DAH7PS enzyme that is involved in secondary metabolism. The absence of the α_{2a} and α_{2b} helices relates directly to the formation of a novel, dimeric, solution-state structure for this type II DAH7PSs with altered allosteric functionality, in-line with the enzyme's role within secondary metabolism. The structure and properties of *Pae*DAH7PS^{PA1901}, in combination with those recently reported for *Pae*DAH7PS^{PA2843} [33], indicate that the evolutionary trajectory for the type II DAH7PSs may have diverged to deliver type II enzymes that function within either primary or secondary metabolism. The distinct structural and functional properties of *Pae*DAH7PS^{PA1901}, in combination with sequence characteristics of the type II DAH7PSs, suggest that the type II DAH7PSs be further classified into two groups noted here as type II_A and type II_B.

Experimental procedures

Sequence clustering analysis

Sequences of type II DAH7Ps were extracted from the Pfam database [40] (seed sequences of DAHP synthetase II - PF01474), aligned with Clustal Ω [41] and submitted to jackhmmer (part of the HMMER web server [42]) to create a hidden Markov model (HMM) profile used to scan, using an E-value sequence cut-off of 10^{-15} , four different databases: UniProt [43], RefSeq [44], Pfamseq [40] and NR [44]. The E-value cut-off was carefully chosen to ensure that no type I sequences were retrieved but all type II sequences would be. This was possible due to the low sequence similarity between type I and II DAH7PSs, which is much lower compared with the sequence similarity between the two identified clusters of type II DAH7PSs. Therefore, selecting an E-value cut-off at the limit of inclusion of type I sequences enabled us to retrieve all the type II sequences available in those databases, at the time of the analysis. For each database, the scan was iterated until convergence (i.e., no new sequences identified) and the results grouped together and duplicates removed using Jalview [45], leading to 2678 non-redundant type II DAH7PS sequences which were used as an input for the clustering method implemented in CLANS [46]. Specifically, after an all-against-all BLAST search of the sequences, a force-directed pairwise similarities clustering algorithm was run for more than 500 iteration cycles at a *P*-value of 10^{-15} .

Protein expression and purification

The ORF encoding *Pae*DAH7PS^{PA1901} (EC 2.5.1.54) was amplified from *P. aeruginosa* PAO1 gDNA using the PCR. The resultant PCR product was cloned into the expression vector pET-28a(+) and engineered to incorporate an N-terminal tobacco etch virus (TEV) protease-cleavable His₆ purification tag. The complete plasmid was

sequence-verified (MacroGen), transformed into *Escherichia coli* BL21*(DE3) cells and co-expressed with the chaperonins pGroES and pGroEL. Expression was achieved following the addition of 1 mM IPTG and subsequent incubation at 23°C for 16 h. Cells were harvested by centrifugation (12000 g, 15 min). Cell lysis was achieved in lysis buffer (10 mM bis-tris propane pH 8.0, 200 mM KCl, 1 mM tris(2-carboxyethyl)phosphine hydrochloride, 200 μM PEP, 10 mM imidazole) by sonication (4 × 5-min cycles at 80% power). Cellular DNA was degraded by the addition of benzonase before the removal of cellular debris by centrifugation (40000 g, 30 min). Purification was carried out using Co²⁺ affinity chromatography, incubation with TEV protease (4°C, 3 h), and size-exclusion chromatography. In brief, the soluble fraction of the cell lysate (containing PaeDAH7PS^{PA1901}) was loaded on to a talon trap column pre-equilibrated with lysis buffer. Contaminating *E. coli* proteins were washed through the column before isocratic elution of PaeDAH7PS^{PA1901} in buffer containing 10 mM bis-tris propane pH 8.0, 200 mM KCl, 1 mM tris(2-carboxyethyl)phosphine hydrochloride, 200 μM PEP, 100 mM imidazole. Protein samples were diluted (1:1) with lysis buffer immediately after elution from the column. The His₆ purification tag was cleaved by incubation with TEV protease (2 mg, 4°C, 3 h) before the cleaved tag was removed from the protein sample by a second round of affinity purification. Protein samples were concentrated and loaded on to a Hiload™ 26/30 Superdex™ 200 column pre-equilibrated with buffer containing 10 mM bis-tris propane pH 8.0, 200 mM KCl, 1 mM tris(2-carboxyethyl)phosphine hydrochloride, 200 μM PEP. Protein concentrations were determined using a Nanodrop ND-1000 spectrophotometer, at 280 nm, using the molar extinction coefficient (54430 M⁻¹.cm⁻¹) calculated for the protein using ProtParam. Purified protein samples were flash frozen in liquid nitrogen and stored at -80°C.

MS

The molecular weight of PaeDAH7PS^{PA1901} was determined by ESI MS (Bruker maXis 3G). Protein samples were dialysed into Milli-Q water and diluted to a concentration of 0.3 mg.ml⁻¹ prior to analysis. The molecular mass of a single chain of PaeDAH7PS^{PA1901} was found to be 44470 Da compared with the calculated theoretical mass of 44468 Da (ProtParam).

Enzyme kinetic assays

The activity of PaeDAH7PS^{PA1901} was monitored over a range of temperatures (from 35 to 50°C) and a range of pHs (pH 6.5–8.5) based on methods previously described [26] using a Varian Cary 300 UV-Vis spectrophotometer. Metal ion dependency was determined by monitoring the activity of PaeDAH7PS^{PA1901} in the presence of 100 μM of various divalent metal cations. The enzyme was pre-treated with EDTA (0.5 mM, 2 h) to remove background metal ions before being buffer-exchanged into assay buffer that had been pre-treated with Chelex (Bio-Rad). PEP (Sigma) and E4P (Sigma) concentrations were held constant at 150 μM, except when determining the respective K_M values, determined by monitoring the activity of PaeDAH7PS^{PA1901} in the presence of 10–200 μM (E4P) or 10–400 μM (PEP) of the substrate for which K_M was being measured. For the inhibition studies, stock solutions of either Trp, Tyr or Phe were prepared in ultrapure water. Stock solutions of phenazine or PCA were prepared in DMSO and activity was compared with controls where phenazine or PCA was substituted for an equivalent amount of DMSO. All reactions were carried out in the presence of 100 μM Co²⁺, except when determining metal ion preference, and the reaction was initiated by the addition of purified PaeDAH7PS^{PA1901}. Initial reaction rates were determined using a least-squares fit of the data.

Analytical ultracentrifugation

Sedimentation velocity experiments were performed in a Beckman Coulter Model XL-I analytical ultracentrifuge equipped with UV/Vis scanning optics. Reference buffer solution (50 mM bis-tris propane, pH 7.5, 200 mM KCl, 100 μM cobalt chloride, 200 μM PEP) and sample solutions (including reference buffer solution with PaeDAH7PS^{PA1901} at three concentrations: 0.34 mg.ml⁻¹ (8 μM), 1.0 mg.ml⁻¹ (23 μM), and 1.35 mg.ml⁻¹ (30 μM)) were loaded into 12-mm double-sector cells with standard Epon 2-channel centerpieces and sapphire windows. For the two higher concentrations (23 and 30 μM), cells were mounted in an eight-hole An-50 Ti rotor and centrifuged at 50000 rpm at 20°C, with absorbance measurements at a wavelength of 295 nm (collected in intensity mode) recorded over a radial position range of 5.8–7.3 cm within the cells taken at sediment boundary intervals of 0.003 cm. In order to gain a more optimal signal-to-noise ratio for the lowest concentration (8 μM) and buffer without protein present, cells were mounted in a four-hole An-60 Ti rotor and centrifuged at 40000 rpm at 20°C, with absorbance measurements at a wavelength of 240 nm (collected in intensity mode) recorded over a radial position range of 5.8–7.3 cm within the cells taken at sediment boundary intervals of 0.003 cm. Further sedimentation velocity experiments, utilising protein at 17 μM, in the presence or absence of 200 μM of either PYO, Phe, Tyr or Trp, were carried out using an

eight-hole An-50 Ti rotor and centrifuged at 35000 rpm at 20°C, with absorbance measurements at a wavelength of 290 nm recorded over a radial position range of 5.8–7.3 cm within the cell taken at sediment boundary intervals of 0.003 cm. Buffer density (1.0129 g/ml) and buffer viscosity (1.050 cP) were experimentally measured with an Anton Paar DMA4100M density meter and Anton Paar Lovis 2000 ME microviscometer respectively. The 2DSA-Monte Carlo, van Holde-Weischet, and Discrete Model Genetic Algorithm (DMGA) analyses were performed using UltraScan III [47–50]. Bead modelling and hydrodynamic calculations were performed using UltraScan Solution Modeller (US-SOMO) [51,52].

Small angle X-ray scattering data collection and analysis

Size-exclusion chromatography coupled small angle X-ray scattering (SEC-SAXS) data were collected at the SAXS/WAXS beamline at the Australian Synchrotron [53] using a sheath flow sample environment [54] at 12 keV (1.0332 Å), using a detector distance of 1600 mm, and at a temperature of 293 K. Data were collected immediately after elution from a Superdex S200 (5 × 150 mm) column at a flow rate of 0.2 ml.min⁻¹ [55]. Samples were loaded on to the column at protein concentrations of 8.0, 5.0 and 1.0 mg.ml⁻¹ in buffer containing 50 mM bis-tris propane pH 7.5, 100 μM cobalt chloride, 200 μM PEP, 5% glycerol.

Data were processed using the reduction software *ScatterBrain* 2.83, developed at the Australian Synchrotron. Scattering intensity (*I*) was plotted versus *q*, as a log-linear plot, and analysed using the *ATSAS* package [56]. Deconvolution of the data was achieved using the HPLC module of the *SOMO* package [52,57] by fitting two pure Gaussian functions to each SEC-SAXS dataset. *GASBOR* [58] was used to generate *ab initio* dummy residue models from the *P(r)* obtained from the deconvoluted data for peaks A and B, which were overlaid with the crystal structure of *Pae*DAH7PS^{PA1901} (Protein Data Bank (PDB): 6BMC).

Crystallography and structure determination

Protein crystals were prepared, by microbatch crystallisation [59], by mixing equal volumes of purified protein (final protein concentration 3–5 mg.ml⁻¹ (67–112 μM)) with reservoir solution (0.2 M sodium fluoride, 1 mM cobalt chloride, 1 mM PEP, 18% PEG 3350) and incubating at 278 K for 1–2 days. Crystals were flash frozen at 110 K in cryoprotectant containing 25% glycerol and mother liquor. X-ray diffraction data were collected at the Australian Synchrotron using the MX2 beamline [60], equipped with an Eiger 16M detector, at a wavelength of 0.9536 Å. Diffraction data were processed using XDS [61] and AIMLESS [62], and the structure of *Pae*DAH7PS^{PA1901} was solved by molecular replacement (MOLREP) [63] using a single chain of *Pae*DAH7PS^{PA2843} (PDB: 5UXM) [33] as the search model. All ligands and waters were removed from the search model prior to molecular replacement, as were residues corresponding to the inserted helices α_{2a} and α_{2b}. The sequence identity between the search model and the target protein was 43%. The model was built using COOT [64] and refined with REFMAC [65].

Interface analysis

PISA [66] was used to visualise and examine the residues involved in interface formation. LSQKAB [67] was used to superpose and compare the structures.

PDB accession codes

Atomic co-ordinates and structure factors for the structure described in this work have been deposited in the PDB with the accession code 6BMC.

Results and discussion

Clustering of type II DAH7PS sequences reveals an uncharacterised subgroup of type II enzymes

Clustering of type II DAH7PSs, based on pairwise sequence similarity, enables the identification of two main clusters of sequences presenting high intra- and low inter-cluster sequence similarity (Figure 2). The main cluster contains sequences corresponding to full-length type II DAH7PSs (including *Pae*DAH7PS^{PA2843}, *Mtu*DAH7PS and *Cgl*DAH7PS) that contain both an N-terminal extension and the α_{2a} and α_{2b} inserted helices. However, a second distinct group of sequences, which are distant from the main cluster, is also evident. Sequences from this second grouping (of which *Pae*DAH7PS^{PA1901} is a member) are shorter in sequence length, relative to those found in the main type II DAH7PS cluster, due to the predicted omission of the sequence corresponding to the α_{2a} and α_{2b} helices. Although there is high sequence homology among members of each subgrouping (for example, *Pae*DAH7PS^{PA2843}

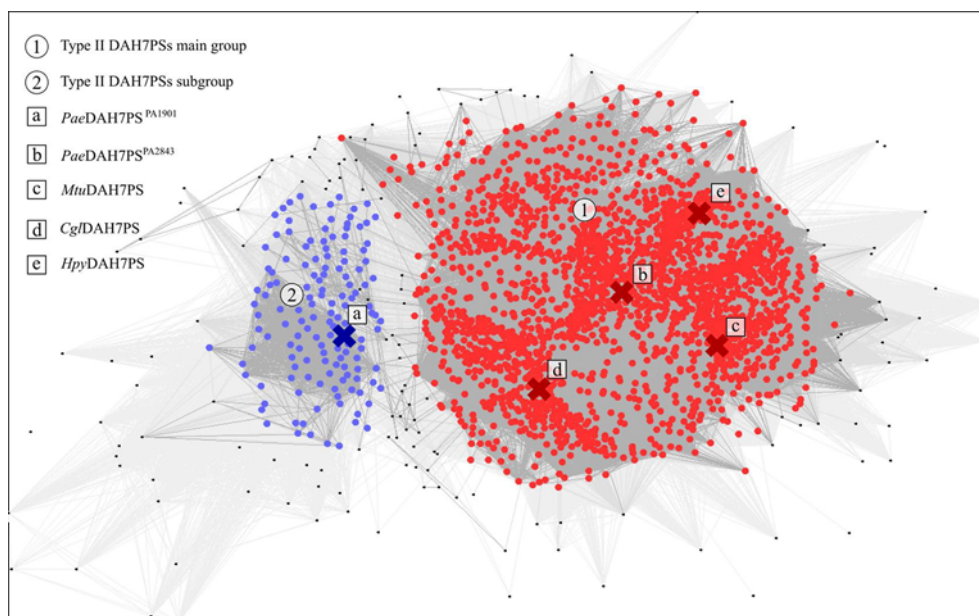


Figure 2. CLANS clustering analysis of type II DAH7PS sequences reveals two distinct groups of type II DAH7PSs

Each dot represents a type II DAH7PS sequence. The main group of type II DAH7PSs (1) is indicated by the red dots. The second group of type II DAH7PSs (2) is indicated by the blue dots. Lines connecting the dots indicate the sequence similarity relationship at the BLAST P -value cut-off of 10^{-50} , the darker the colour, the higher the sequence similarity. Crosses marked (a–e) correspond to the sequences of *PaeDAH7PS*^{PA1901}, *PaeDAH7PS*^{PA2843}, *MtuDAH7PS*, *CglDAH7PS* and *Helicobacter pylori* DAH7PS (*HpyDAH7PS*) respectively.

Table 1 Kinetic parameters determined for *PaeDAH7PS*^{PA1901} at pH 7.5, 37°C and comparison with those previously reported for either *PaeDAH7PS*^{PA2843} or *MtuDAH7PS*

Enzyme	Metal ion	k_{cat} (s ⁻¹)	K_M^{E4P} (μM)	K_M^{PEP} (μM)	k_{cat}/K_M^{E4P} (s ⁻¹ ·μM ⁻¹)	k_{cat}/K_M^{PEP} (s ⁻¹ ·μM ⁻¹)
<i>PaeDAH7PS</i> ^{PA1901}	Co ²⁺	19.8 ± 0.4	16 ± 3	17 ± 1	1.24 ± 0.15	1.16 ± 0.10
<i>PaeDAH7PS</i> ^{PA2843} [33]	Co ²⁺	40.0 ± 0.7	18 ± 1	28 ± 1	2.20 ± 0.20	1.40 ± 0.10
<i>MtuDAH7PS</i> [26]	Mn ²⁺	3.1 ± 0.1	25 ± 1	37 ± 3	0.12 ± 0.10	0.08 ± 0.01

and *MtuDAH7PS* share 50% sequence identity and 67% sequence similarity), a comparison between sequences from the main cluster with those from the subgroup reveals increased sequence diversity between the two type II DAH7PS groups. For example, *PaeDAH7PS*^{PA1901} and *MtuDAH7PS* share only 38.5% sequence identity and 50.0% sequence similarity, and *PaeDAH7PS*^{PA1901} and *PaeDAH7PS*^{PA2843} share 38.4% sequence identity and 52.0% sequence similarity. Does this difference in sequence characteristics translate to altered structural and/or functional properties for this second uncharacterised group of type II DAH7PSs, analogous to those observed for the type Iα compared with type Iβ DAH7PSs? To address this question, we sought full characterisation of *PaeDAH7PS*^{PA1901}.

***PaeDAH7PS*^{PA1901} is insensitive to aromatic amino acids or PCA**

The purified recombinant *PaeDAH7PS*^{PA1901} was found to be catalytically active over a range of temperatures between 35 and 50°C and over a range of pH between pH 6.5 and 7.5 (Supplementary Figure S2), in contrast with *PaeDAH7PS*^{PA2843} where maximal activity is observed over a narrow range of temperatures and pH [33]. Maximal *PaeDAH7PS*^{PA1901} activity was observed at pH 7.5 and 45°C. Metal ion preference was investigated by monitoring the activity of *PaeDAH7PS*^{PA1901} in the presence of various divalent metal cations, and it was found that Mn²⁺ was most the activating (Figure 3A). Subsequent assays were carried out at pH 7.5, 37°C in the presence of Co²⁺ in order to provide a comparison with *PaeDAH7PS*^{PA2843}, which exhibits maximal activity under these conditions [33]. Apparent K_M values for *PaeDAH7PS*^{PA1901} for PEP and E4P were determined to be 17 ± 1 and 16 ± 3 μM respectively (Table 1). The Michaelis constants are in-line with other characterised type II DAH7PSs [26,33,39,68], including

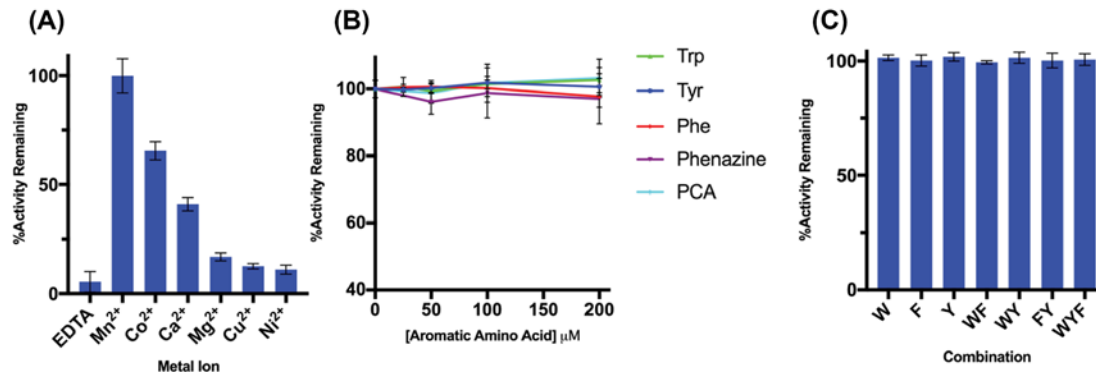


Figure 3. Activity of *PaeDAH7PS*^{PA1901}

(A) In the presence of 100 μM of various divalent metal cations or 100 μM of EDTA. (B) In the presence of single aromatic amino acids or secondary metabolites (Trp, green; Tyr, blue; Phe, red; phenazine, purple; PCA, cyan) or (C) binary and ternary combinations of aromatic amino acids. Each single letter code corresponds to 100 μM of the corresponding amino acid. PEP and E4P concentrations were held constant for all measurements at 150 μM each. Error bars represent the S.D. of triplicate measurements.

PaeDAH7PS^{PA2843}, and the turnover number, k_{cat} , for *PaeDAH7PS*^{PA1901} was determined to be $19.8 \pm 0.4 \text{ s}^{-1}$.

The activity of *PaeDAH7PS*^{PA1901} was monitored in the presence of increasing concentrations of the aromatic amino acids Trp, Tyr, Phe or the secondary metabolites phenazine or PCA. At concentrations up to 200 μM Trp, Tyr, Phe, phenazine or PCA, *PaeDAH7PS*^{PA1901} activity was found to be comparable with that observed in the absence of aromatic amino acids or secondary metabolites, analogous to the allosteric behaviour of the unregulated type I β DAH7PSs [69] (Figure 3B,C). Combinations of aromatic amino acids appear to have no inhibitory effect on *PaeDAH7PS*^{PA1901} activity similar to that observed in the absence of aromatic amino acids (Supplementary Figure S3). The observed absence of allosteric sensitivity in *PaeDAH7PS*^{PA1901} is in contrast with *MtuDAH7PS* or *PaeDAH7PS*^{PA2843} where allosteric inhibition was observed under the same conditions that were used to evaluate the allosteric properties of *PaeDAH7PS*^{PA1901}. In particular, in *MtuDAH7PS*, any binary or ternary combination of aromatic amino acids that includes Trp acts to synergistically inhibit the enzyme [34–36] or, in *PaeDAH7PS*^{PA2843}, sensitivity to Trp alone was observed, but this sensitivity was diminished in comparison with that observed for *MtuDAH7PS* [33].

The crystal structure of *PaeDAH7PS*^{PA1901} reveals novel quaternary assembly

The crystal structure of *PaeDAH7PS*^{PA1901} (phzC) was solved (resolution 2.70 \AA , $R_{\text{free}} = 0.280$) in complex with the substrate PEP and a Co^{2+} ion, with attached water molecule, bound at the active site, revealing for the first time the structure of a short-form type II DAH7PS that is involved in secondary (here phenazine) metabolism. *PaeDAH7PS*^{PA1901} crystallised in the space group $C222_1$, with two DAH7PS chains present in the asymmetric unit. Application of a two-fold crystallographic symmetry operation results in the assembly of a homotetrameric species, which comprises both a major and minor interfaces. Chain A residues 119–123, 172–177 and 389–405, and chain B residues 121–123, 170–177 and 389–405 are not resolved in this structure and were therefore not included in the final model (Figure 4). Data collection and refinement statistics are shown in Table 2.

As with all DAH7PS structures reported to date [22–33], *PaeDAH7PS*^{PA1901} features a core $(\beta/\alpha)_8$ -barrel fold, with an N-terminal extension to the core catalytic domain consistent with its membership of the type II DAH7PS family (Figure 4). Residues 1–59 form an N-terminal extension to the barrel, providing additional helices α_{0a} , α_{0b} and α_{0c} , with strong structural homology to the equivalent helices in other structurally characterised type II DAH7PSs, in particular *PaeDAH7PS*^{PA2843} [33]. Residues 167–181 form loop $\alpha_2\beta_3$, which lacks the inserted helices α_{2a} and α_{2b} as observed in both *MtuDAH7PS* and *PaeDAH7PS*^{PA2843} [26,33].

The active site for *PaeDAH7PS*^{PA1901} is located at the C-terminal end of the core $(\beta\alpha)_8$ catalytic barrel and is comparable with that observed among the type II DAH7PSs in terms of residue identity. The PEP phosphate group is co-ordinated by atoms Glu217_N, Arg218_{NH1}, Arg271_{NE}, Arg271_{NH2} and Lys240_{NZ} whereas the carboxylate group of PEP is co-ordinated by atoms Arg106_{NH1} and Lys240_{NZ} (Figure 5 and Supplementary Figure S4).

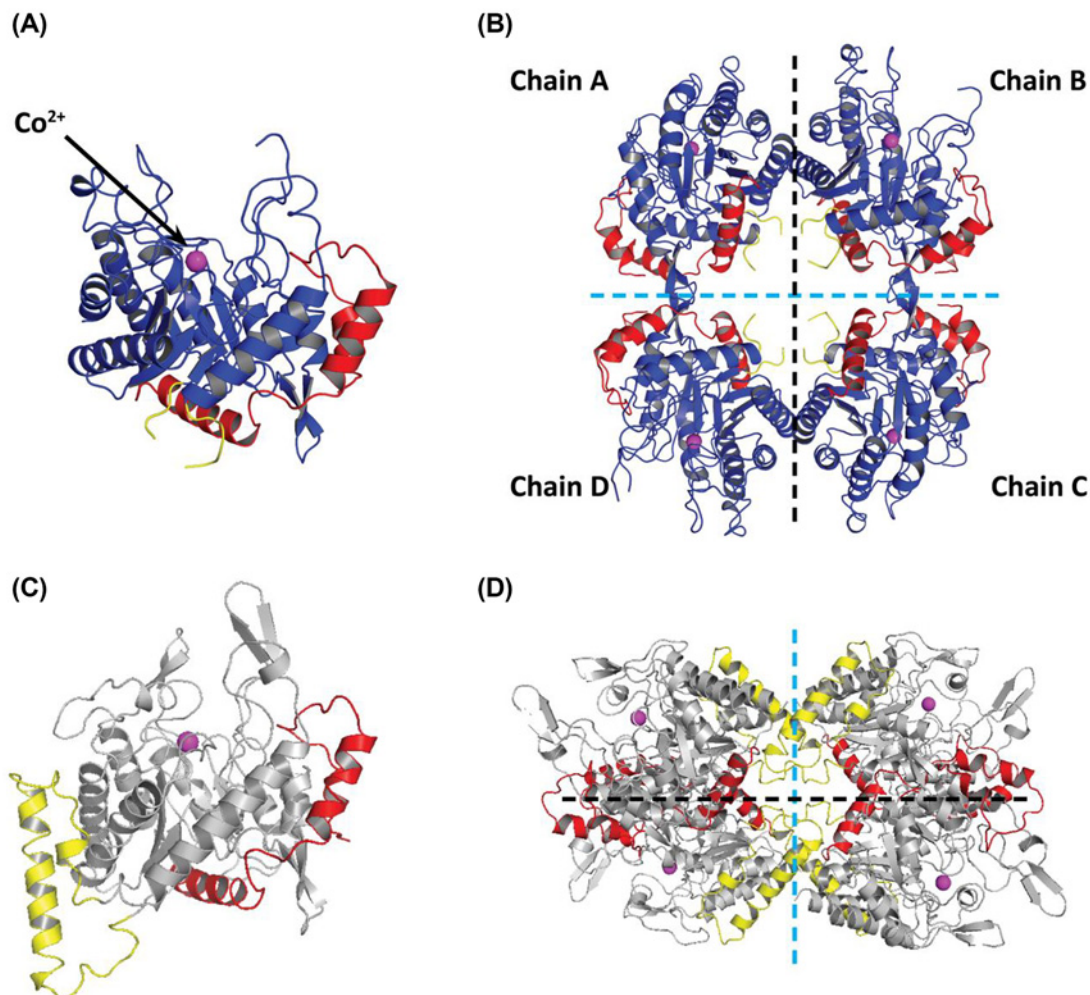


Figure 4. Comparison among the crystal structures of *PaeDAH7PS*^{PA1901} and *PaeDAH7PS*^{PA2843}

(A) A single chain of *PaeDAH7PS*^{PA1901} (PDB: 6BMC). The core ($\beta\alpha$)₈ catalytic domain is shown in blue, the N-terminal extension (residues 1–59) is shown in red, the short extension to loop $\alpha_2\beta_3$ (residues 167–181) is shown in yellow (residues 172–177 are unresolved in this structure), and the active site Co^{2+} ion is shown in magenta. (B) The homotetrameric structure of *PaeDAH7PS*^{PA1901} as observed in the crystal structure. The major interface (indicated by the black dashed line) is formed primarily by residues located on helices α_1 and α_8 . The minor interface (indicated by the cyan dashed line) is formed by residues located on loop $\alpha_3\beta_4$. (C) A single chain of *PaeDAH7PS*^{PA2843} (PDB: 5UXM) shown in grey with the N-terminal extension (residues 1–59) shown in red, the inserted helices α_{2a} and α_{2b} (residues 179–225) shown in yellow and the active site Co^{2+} ion shown in magenta. (D) The homotetrameric structure of *PaeDAH7PS*^{PA2843}. The major interface (indicated by the black dashed line) is formed primarily by residues located on helix α_2 . The minor interface (indicated by the cyan dashed line) is formed by residues primarily located on helix α_{2b} .

⁹⁵The major interface is formed by residues primarily located on helices α_4 and α_5 (and β -strands β_{5a} and β_{5b}) for the type I α DAH7PSs [22,30,70] or by residues primarily located on helices α_6 and α_7 for the type I β DAH7PSs [25,31] or by residues primarily located on helices α_{0c} and α_2 for the type II DAH7PSs [26,32,33]. Based on inspection of the protein sequence, and comparison with that of *PaeDAH7PS*^{PA2843} or *MtuDAH7PS*, we naïvely anticipated that the interface involving helices α_{0c} and α_2 in *PaeDAH7PS*^{PA2843} or *MtuDAH7PS* would be preserved in *PaeDAH7PS*^{PA1901}. Surprisingly, this is not the case (Figure 6). The major interface for *PaeDAH7PS*^{PA1901} is instead formed through interactions between chains A and B (or chains C and D) primarily involving residues located on helices α_1 and α_8 , assembling a completely distinct major interface compared with that observed for any of the DAH7PSs characterised to date. For *PaeDAH7PS*^{PA1901}, a total of 22 residues from chain A and 23 residues from chain B are involved in the formation of this major interface, burying 840 Å² (5.5%) or 819 Å² (5.3%) of the surface area of each chain respectively, comparable with the surface area involved in the formation of the major interface observed

Table 2 Data collection and refinement statistics for the crystal structure determined for *PaeDAH7PS*^{PA1901}

PDB code	6BMC
Space group	C222 ₁
Data collection and processing statistics	
Unit cell dimensions (Å)	
<i>a</i>	53.8
<i>b</i>	169.7
<i>c</i>	170.6
Resolution range (outer shell) (Å)	49.08–2.70 (2.83–2.70)
Wavelength (Å)	0.9536
Unique reflections	21959 (2879)
Completeness (%)	99.9 (100.0)
Redundancy	9.4 (10.0)
<i>I</i> /σ	10.6 (0.8)
<i>R</i> _{merge}	0.128 (2.619)
<i>R</i> _{pim}	0.064 (1.280)
<i>CC</i> _{1/2}	0.999 (0.461)
Refinement resolution (Å)	2.70 (2.83–2.70)
Refinement statistics	
<i>R</i> factor	0.23
<i>R</i> _{free} (5 %)	0.28
Number of non-hydrogen atoms	
Protein atoms	5667
Water atoms	13
Ligand atoms	22
Wilson <i>B</i> value (Å ²)	73.7
Mean <i>B</i> value (Å ²)	
All atoms	88.1
Protein atoms	91.2
Ligand atoms	65.1
Water atoms	83.1
RMSDs from ideality	
Bond lengths (Å)	0.004
Bond angles (°)	0.800
Residues in the most favoured region of Ramachandran plot (%)	95.45
Residues in the forbidden region of Ramachandran plot (%)	0.27 ¹

¹Cys³⁷² from both chains A and B.

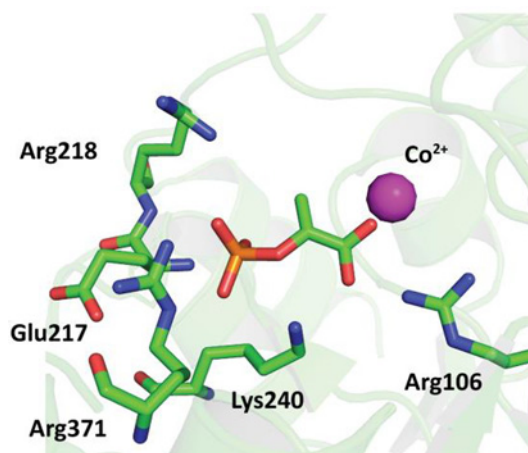


Figure 5. The active site of *PaeDAH7PS*^{PA1901} (PDB: 6BMC) showing the binding mode of the substrate PEP

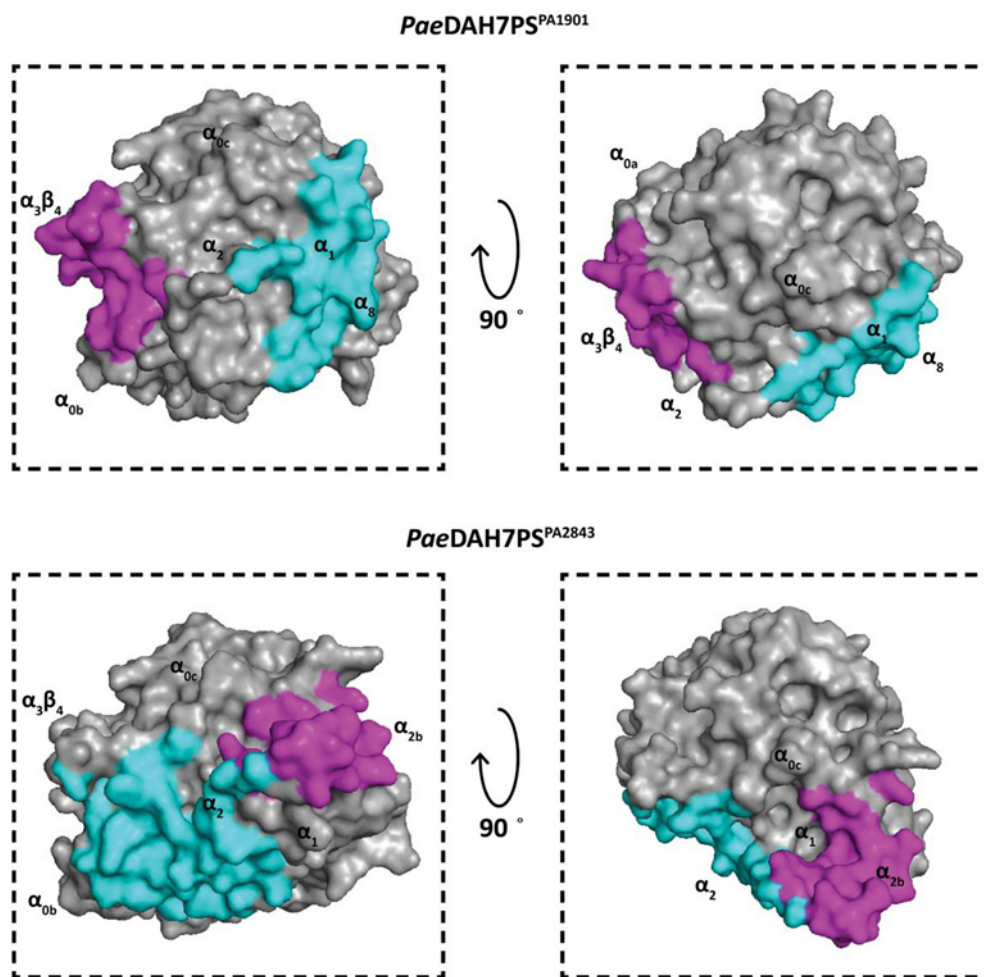


Figure 6. Comparison between the regions of the $(\beta/\alpha)_8$ -barrel involved in the formation of the oligomeric interfaces in *Pae*DAH7PS^{PA1901} (PDB: 6BMC) and *Pae*DAH7PS^{PA2843} (PDB: 5UXM)

For all figures, surface area involved in the formation of the major interface is highlighted in cyan and surface area involved in the formation of the minor interface is highlighted in magenta. The upper pair of figures is shown in the same orientation relative to the corresponding figure in the lower pair of figures. The approximate locations of key structural elements are indicated.

for either *Pae*DAH7PS^{PA2843} or *Mtu*DAH7PS [26,32,33]. Four equivalent pairs of salt bridges were identified to form between chain A atoms Glu77_OE1 (and Glu77_OE2), Glu87_OE1 (and Glu87_OE2), Arg94_NH1 (and Arg94_NH2), Arg103_NH1 (and Arg103_NH2) and chain B atoms Arg94_NH2, Arg103_NH2, Glu77_OE1, and Glu87_OE1 respectively (Figure 7). A pair of equivalent hydrophobic contacts are made between chain A residue Leu⁹⁵ and chain B residues Pro³⁴⁸ and Val⁸⁵ (and *vice versa*), and between chain A residue Ala⁹² and chain B residue Leu⁸⁸ (and *vice versa*). Hydrophobic contacts are also found between chain A residue Leu⁸⁸ and two residues from chain B: Leu and Leu⁸⁸ (and *vice versa*). Further hydrophobic contacts are found between chain A residues Trp³⁷⁹, Leu³⁸² and chain B residues Met³⁸⁶ and Leu³⁸² respectively.

Unexpectedly, despite the absence of the α_{2a} and α_{2b} helices in *Pae*DAH7PS^{PA1901}, which are associated with the formation of the minor interface in either *Pae*DAH7PS^{PA2843} or *Mtu*DAH7PS, an alternative interaction to deliver a homotetrameric species is observed in *Pae*DAH7PS^{PA1901} (Supplementary Table S3). This alternative minor interface for *Pae*DAH7PS^{PA1901} is formed through interaction between chains A and D (or chains B and C) by residues located primarily on loop $\alpha_3\beta_4$ and is distinctive from the minor interface observed for all characterised DAH7PSs (both type I and II). In comparison, the minor interface is formed by residues primarily located on helix α_{0a} for the type I α DAH7PSs [22,30,70] or by residues located on helices α_4 and α_5 for the type I β DAH7PSs [25,31] or by residues primarily located on helices α_{2a} and α_{2b} for the type II DAH7PSs [26,32,33]. For *Pae*DAH7PS^{PA1901}, a total of 20 residues each from chain A or D are involved in the formation of this minor interface with a buried surface area of 598

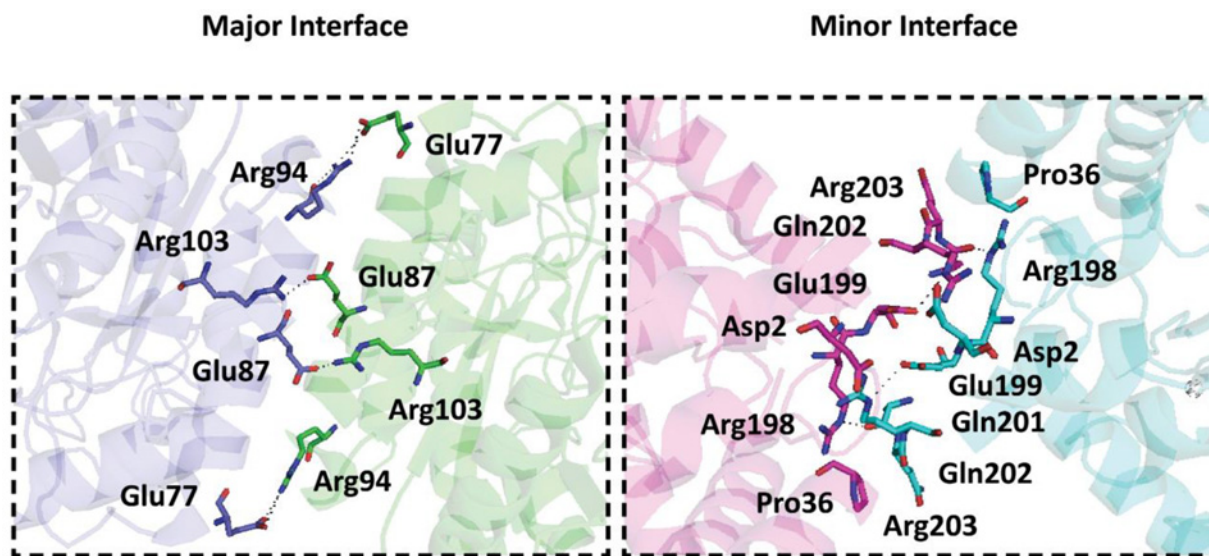


Figure 7. Charged residues involved in the formation of both the major and minor interface of *PaeDAH7PS*^{PA1901}

For the major interface, chain A is shown in blue and chain B is shown in green. For the minor interface, chain A is shown in magenta and chain D is shown in cyan. For both figures, polar contacts are indicated by the black dashed lines.

\AA^2 (4.0%) or 602\AA^2 (4.0%) respectively. A pair of salt bridges is formed between chain A atom Asp2_OD1 and chain D atom Arg203_NH2 (likewise for chain A atom Arg203_NH2 and chain D atom Asp2_OD1) as well as a hydrogen bond between chain A atom Arg198_NE and chain D atom Glu201_O (likewise for chain A atom Glu201_O and chain D atom Arg198_NE). In addition, a limited suite of hydrophobic contacts is found between methylene groups of Gln²⁰² and Arg¹⁹⁹ in chain A and Pro³⁶ and Arg²⁰³ in chain D (and *vice versa*).

For *MtuDAH7PS*, three distinct aromatic amino acid allosteric binding sites exist that are each selective for either Trp, Tyr or Phe. The Phe and Trp sites are located at the oligomeric interfaces and are intimately associated with the formation of the quaternary assembly [34,36,71]. In comparison, for *PaeDAH7PS*^{PA2843} a single allosteric binding site exists at the tetramer interface that is sensitive for Trp [33] and structurally comparable with the Trp site of *MtuDAH7PS*. For *PaeDAH7PS*^{PA1901}, the alternative oligomeric interfaces and subsequent formation of a significantly different quaternary assembly, relative to either *PaeDAH7PS*^{PA2843} or *MtuDAH7PS*, disrupts completely the formation of any aromatic amino acid allosteric binding sites that are comparable with those observed for either *PaeDAH7PS*^{PA2843} or *MtuDAH7PS*. Consistent with this is the observation made during functional characterisation that *PaeDAH7PS*^{PA1901} is insensitive to allosteric regulation by aromatic amino acids, confirming that *PaeDAH7PS*^{PA1901} functions primarily within secondary metabolism.

Solution-state structure of *PaeDAH7PS*^{PA1901}

SEC-SAXS data were collected using three different starting protein concentrations: 1.0 , 5.0 and 8.0 mg.ml^{-1} (22 – $180 \text{ }\mu\text{M}$) to investigate the solution-state structure of *PaeDAH7PS*^{PA1901} and the concentration dependency of quaternary structure (Figure 8 and Table 3, Supplementary Figure S5 and Tables S1 and S2).

For the SAXS data collected using an injection concentration of 8.0 mg.ml^{-1} ($180 \text{ }\mu\text{M}$), *PaeDAH7PS*^{PA1901} eluted as a single peak with a trailing back edge, indicating polydispersity in the sample. The scattering data were deconvoluted using the HPLC module of the SOMO package through the fitting of Gaussian functions to the SEC-SAXS data [52,55,57]. The analysis indicated that there were at least two protein populations contributing to the single elution peak of the SEC-SAXS data. Two pure Gaussian functions were applied to the data, resulting in two distinct scattering profiles; peak A and peak B. Peak A represents the front edge of the elution peak ($R_g = 36.0 \pm 1.2 \text{ \AA}$, $d_{\text{max}} = 114 \text{ \AA}$) while peak B was found to spread across the entire elution peak ($R_g = 33.0 \pm 1.4 \text{ \AA}$, $d_{\text{max}} = 99 \text{ \AA}$). The calculated d_{max} values from the crystal structure of *PaeDAH7PS*^{PA1901} (PDB: 6BMC) for the tetramer, dimer, or monomer are 115.5 , 93.3 , or 62 \AA respectively, with the calculated d_{max} values for peaks A and B more closely resembling that determined from the tetrameric or dimeric crystal structures of *PaeDAH7PS*^{PA1901} respectively. In addition, the calculated R_g values from the crystal structure of *PaeDAH7PS*^{PA1901} for the tetrameric, dimeric, or monomeric species are 39.2 , 29.2 , and 20.9 \AA respectively, with the calculated R_g values for peaks A and B more closely resembling those determined

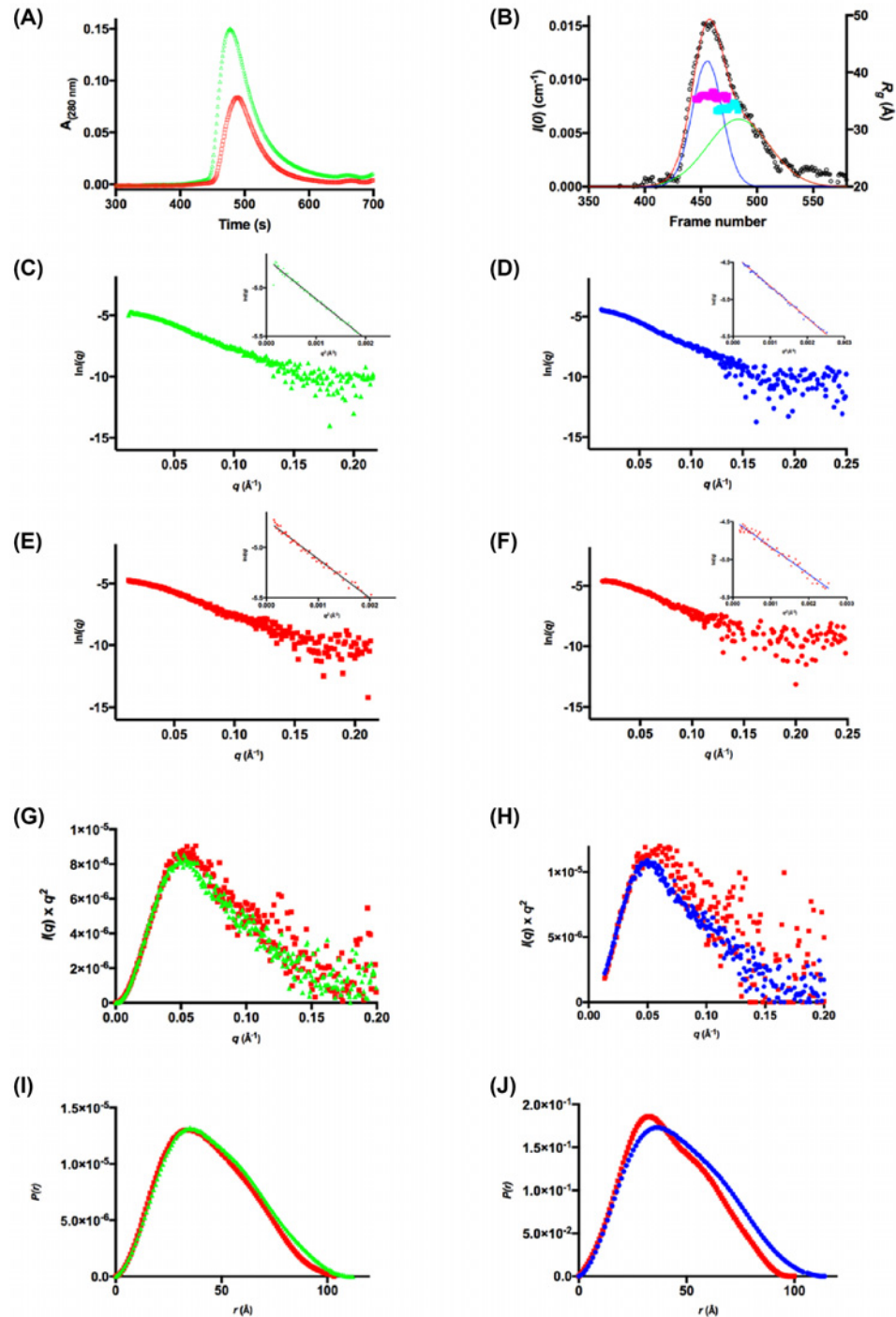


Figure 8. SEC-SAXS analysis for *PaeDAH7PS*^{PA1901}

(A) SEC-SAXS elution profile for two injected enzyme concentrations (5.0 mg.ml⁻¹, red squares and 8.0 mg.ml⁻¹, green triangles). (B) Deconvolution of the SEC-SAXS data indicates two Gaussian components (peak A, blue line and peak B, green line. Sum, red line). The R_g values across each peak are indicated as magenta or cyan squares respectively. (C) the SAXS profile for the non-deconvoluted 8.0 mg.ml⁻¹. (D) The SAXS profile for the deconvoluted 8.0 mg.ml⁻¹ peak A. (E) The SAXS profile for the non-deconvoluted 5.0 mg.ml⁻¹. (F) The SAXS profiles for the deconvoluted 8.0 mg.ml⁻¹ peak B. Guinier plots are inset for frames (C–F). (G) Kratky plots of the non-deconvoluted data in (C,E) (8.0 mg.ml⁻¹, green triangles and 5.0 mg.ml⁻¹, red squares). (H) Kratky plots of the deconvoluted data in (D,F) (peak A, blue circles and peak B, red squares). (I) $P(r)$ plots for the non-deconvoluted data in (C,E) (8.0 mg.ml⁻¹, green triangles and 5.0 mg.ml⁻¹, red squares). (J) $P(r)$ plots for the deconvoluted data in (D,F) (peak A, blue circles and peak B, red squares).

Table 3 SEC-SAXS parameters determined for *Pae*DAH7PS^{PA1901}

	Non-deconvoluted 8.0 mg.ml ⁻¹	8.0 mg.ml ⁻¹ Peak A	8.0 mg.ml ⁻¹ Peak B	1.0 mg.ml ⁻¹
Guinier analysis				
R_g (Å)	34.4 ± 0.6	36.0 ± 1.2	33.0 ± 1.4	34.0 ± 2
$I(0)$ (cm ⁻¹)	0.030 ± 0.000	0.012 ± 0.000	0.012 ± 0.000	0.010 ± 0.001
q_{min}	0.012	0.013	0.013	0.013
Correlation coefficient, r^2	0.978	0.996	0.971	0.960
Pairwise distribution analysis				
R_g (Å)	35.2	36.3	32.8	33.3
$I(0)$ (cm ⁻¹)	0.03	0.01	0.01	0.01
d_{max} (Å)	116.5	114.3	99.0	100.2
V_p (Å ³)	139000	153000	134000	170000
q range (Å ⁻¹)	0.012–0.20	0.013–0.20	0.013–0.20	0.012–0.020
SAXS MoW estimate				
M_W (Da)	Not determined	124000	85000	95000

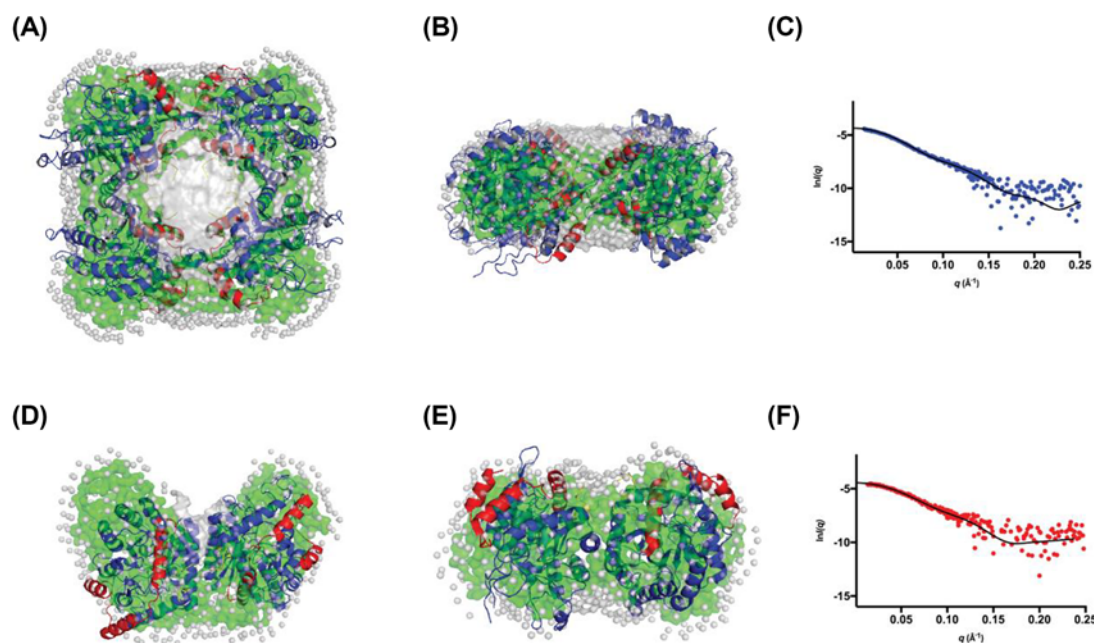


Figure 9. GASBOR modelling results for *Pae*DAH7PS^{PA1901}

(A) GASBOR bead model, generated using the $P(r)$ from peak A (8.0 mg.ml⁻¹), with the tetrameric crystal structure of *Pae*DAH7PS^{PA1901} overlaid. (B) Side view of the model in (A). (C) The fit of the *ab initio* bead model (black line) in (A,B) to the experimental SAXS data (blue circles) from peak A. (D) GASBOR bead model, generated using the $P(r)$ from peak B, with the dimeric crystal structure of *Pae*DAH7PS^{PA1901} overlaid. (E) Side view of the model in (D). For all frames, the core ($\beta\alpha$)₈ catalytic barrel is shown in blue, the N-terminal extension (residues 1–59) is shown in red, the loop $\alpha_2\beta_3$ is shown in yellow. The GASBOR model is represented by the green surface and modelled solvent molecules are represented in grey. (F) The fit of the model (black line) in (D,E) to the experimental SAXS data (red circles) generated from peak B (8.0 mg.ml⁻¹).

from the tetrameric or dimeric crystal structures of *Pae*DAH7PS^{PA1901} respectively. Estimated molecular weights for peaks A and B were calculated using SAXS MoW, which is typically accurate within ±10% [72]. The estimated molecular weights for peaks A and B were 124.5 and 84.6 kDa respectively and are comparable, albeit slightly smaller, with the expected molecular weights for the tetrameric or dimeric *Pae*DAH7PS^{PA1901} of 177.88 and 88.94 kDa respectively.

Ab initio bead models (GASBOR) were generated from the deconvoluted scattering profiles obtained for both peaks A and B to reconstruct the solution-state tetrameric and dimeric species of *Pae*DAH7PS^{PA1901} and to compare the resultant bead models with the oligomeric assemblies observed in the crystal structure (PDB: 6BMC) (Figure 9).

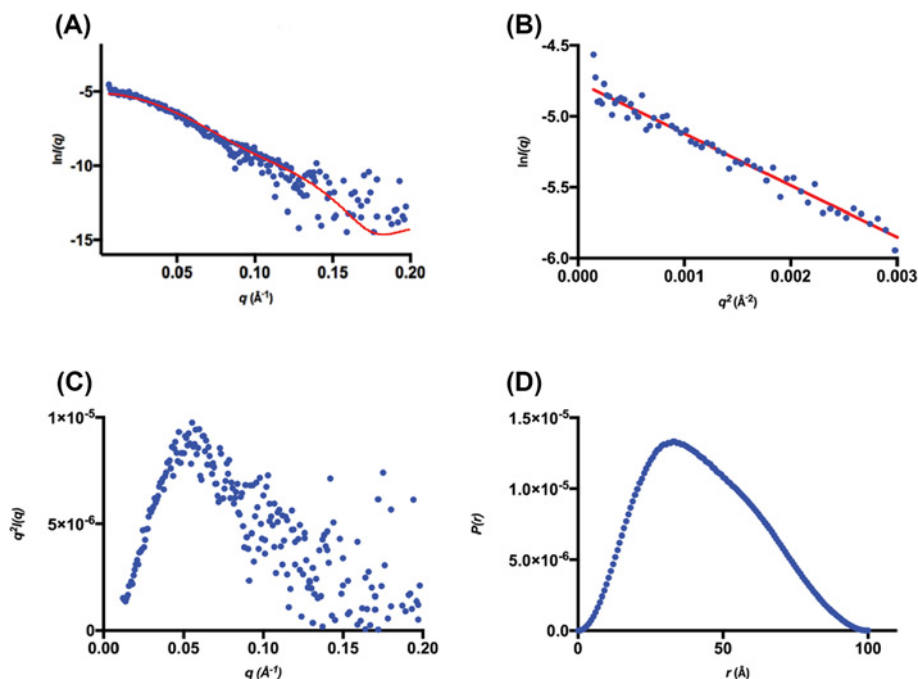


Figure 10. Analysis of SEC-SAXS results obtained for *PaeDAH7PS*^{PA1901}

Using a 1.0 mg.ml⁻¹ injection concentration. (A) $\log I(q)$ compared with q , error bars are indicated in grey, with the theoretical scattering profile calculated from the crystallographic dimer (PDB: 6BMC) overlaid (red line). (B) Guinier plot ($\ln I$) compared with q^2 . (C) Kratky plot ($q^2 \cdot I(q)$) compared with q for the data in (A). (D) $P(r)$ compared with r profiles for the data in (A).

Comparison between the theoretical scattering profiles calculated from the *ab initio* models and the deconvoluted experimental data (Figure 9C,F) suggests that the *ab initio* models are representative of the solution-state tetrameric and dimeric species of *PaeDAH7PS*^{PA1901}, which are remarkably similar to those observed in the crystal structure.

Due to the decreased signal-to-noise ratio for the SEC-SAXS data collected using an injection concentration of 1.0 mg.ml⁻¹ (22 μM), deconvolution of this dataset was not attempted. CRY SOL analysis of the SEC-SAXS data, collected using an injection concentration of 1.0 mg.ml⁻¹, indicates that the enzyme exists primarily in the dimeric form ($\chi^2 = 0.31$ for the fit of the dimeric crystal structure PDB: 6BMC to the experimental data, Figure 10). The d_{\max} value determined from the 1.0 mg.ml⁻¹ SEC-SAXS data of 100.2 Å is consistent with the d_{\max} value determined either from the dimeric crystal structure of *PaeDAH7PS*^{PA1901} (93.3 Å) or for the deconvoluted peak B (99.0 Å). In addition, the SAXS MoW estimated molecular weight of 95.0 kDa from this low concentration SEC-SAXS data is in close agreement, albeit slightly larger, with the value estimated from the deconvoluted peak B (84.6 kDa) and the expected molecular weight for dimeric *PaeDAH7PS*^{PA1901} (88.94 kDa). The SEC-SAXS parameters determined for the data collected using an injection concentration of 1.0 mg.ml⁻¹, in combination with those determined for the deconvoluted 8.0 mg.ml⁻¹ data, show that *PaeDAH7PS*^{PA1901} exists in a concentration-dependent equilibrium that favours the dimeric form on decreasing enzyme concentration.

Analytical ultracentrifugation (AUC) experiments carried out at enzyme concentrations ranging from 0.34 to 1.35 mg.ml⁻¹ (8–30 μM) were used to confirm the oligomeric state of *PaeDAH7PS*^{PA1901} in solution. Analyses of the absorbance data, collected in intensity mode, by van Holde–Weischet analysis reveal half-parabola shaped s-distributions, which shift to the right (Figure 11A) upon increasing protein concentration, suggesting an interacting, reversible system [50]. Non-interacting species between 1–2 S are likely sedimenting buffer components, as illustrated by analysis of buffer without protein present (Figure 11A). 2DSA–Monte Carlo sedimentation coefficient distributions reveal species with sedimentation coefficients between 5.8 and 6.8 S (Figure 11B), consistent with a molecular weight in the range of 70–96 kDa (Supplementary Figure S6), suggesting that at these concentrations, *PaeDAH7PS*^{PA1901} exists predominantly as a homodimer. Species at ~3 S, present in the 8 μM distribution (collected at 240 nm), are likely buffer components that absorb at wavelengths lower than 280 nm, as these species are also present in distributions (also collected at 240 nm) of buffer without protein (data not shown), and to a lesser extent in the 11, 23, and 30 μM samples (Figure 11B). A bead model based on the dimeric crystal structure of *PaeDAH7PS*^{PA1901} (PDB:

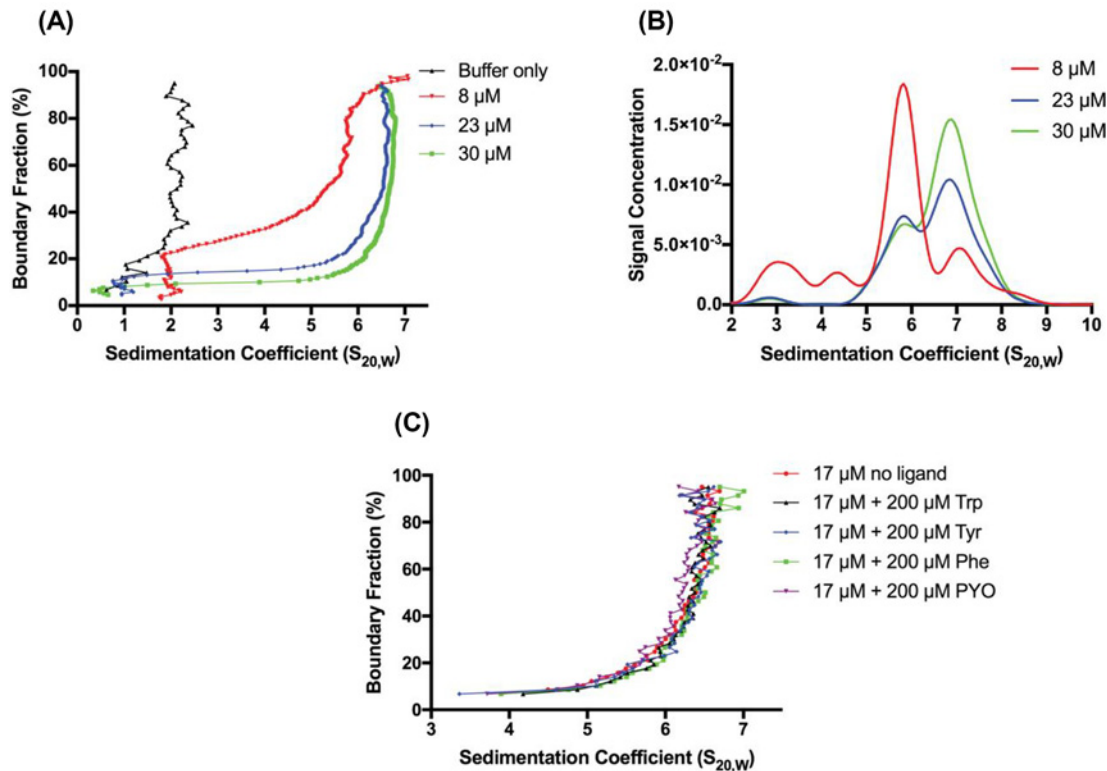


Figure 11. Sedimentation velocity data obtained for *PaeDAH7PS*^{PA1901}

(A) van Holde-Weischet distributions for *PaeDAH7PS*^{PA1901} at three concentrations (8, 23 and 30 μ M) show a shift of the distributions to the right with increasing concentration. (B) Combined $S_{20,w}$ distribution plots from 2DSA-Monte Carlo analysis reveal major species between 5.8 and 6.8 S. (C) van Holde-Weischet analysis of *PaeDAH7PS*^{PA1901} (17 μ M) indicates no significant change in the oligomeric state of the protein in the presence of either 200 μ M of PYO, Phe, Tyr or Trp.

6BMC) was created with US-SOMO and used to calculate a theoretical sedimentation coefficient of 5.5 S, further suggesting that the species observed for *PaeDAH7PS*^{PA1901} is primarily dimeric. Additional sedimentation velocity experiments, carried out in absorbance mode in the presence of 200 μ M of either PYO, Phe, Tyr or Trp, and analysed by van Holde-Weischet analysis, indicate that the presence of either PYO or aromatic amino acids does not influence the oligomeric state of the protein (Figure 11C).

While the formation of a tetrameric species for *PaeDAH7PS*^{PA1901} is observable both in the crystal structure and in solution by SAXS at high injection concentrations (112–180 μ M), the nature of the alternative minor interface (and lack of hydrophobic interactions), in combination with the observation of a primarily dimeric species by AUC at protein concentrations less than 30 μ M, suggests that at physiological concentrations *PaeDAH7PS*^{PA1901} predominantly persists in the dimeric form. The observation of higher-order solution-state species by SEC-SAXS appears to be the consequence of high enzyme concentration.

Evolutionary implications

The structural similarities between the N-terminal extensions (helices α_{0a} , α_{0b} and α_{0c}) found in *PaeDAH7PS*^{PA1901}, *PaeDAH7PS*^{PA2843} or *MtuDAH7PS*, suggest a common origin for this structural element in the type II DAH7PSs. The distinct functionalities of the N-terminal extension within these three enzymes (burying a hydrophobic surface or interface formation for the delivery of allosteric binding sites or combinations thereof), coupled with the physiological roles of these enzymes within primary or secondary metabolism, indicate an evolutionary divergence. The evolutionary trajectory for the type II DAH7PSs appears to have diverged to deliver both an unregulated dimeric group of type II DAH7PSs, suitable for a role within secondary metabolism, and a regulated tetrameric group of type II DAH7PSs that functions within primary metabolism.

For the type II DAH7PSs from *P. aeruginosa*, direct control of enzymatic activity by pathway end products appears largely superfluous as genetic level regulation may be better suited to differentially regulate the expression of multiple DAH7PSs, that function within primary or secondary metabolism, where the presence of aromatic amino acids acts to divert metabolic flux away from primary metabolism and towards the biosynthesis of PCA and its derivatives. Under these conditions, the DAH7PSs that are involved directly within primary metabolism would likely be allosterically inhibited by Trp, Tyr or Phe and thus unavailable to provide chorismate to support the biosynthesis of secondary aromatic metabolites. The presence of *Pae*DAH7PS^{PA1901} within the *phzA–G* biosynthetic cluster allows for the synchronised expression of these proteins required for PCA production. The omission of the α_{2a} and α_{2b} helices in *Pae*DAH7PS^{PA1901}, and subsequent insensitivity to allosteric inhibition by Trp, Tyr or Phe, allows for the continued production of chorismate under conditions of high aromatic amino acids, consistent with the alternative, dimeric solution-state structure observed for *Pae*DAH7PS^{PA1901}.

Conclusion

The structure of *Pae*DAH7PS^{PA1901} further highlights the complex evolutionary trajectory for the type II DAH7PSs that has delivered type II enzymes which exhibit a diverse range of quaternary assemblies, and associated allosteric functionalities, required to support the efficient production of chorismate within either primary or secondary metabolism. *Pae*DAH7PS^{PA1901} adopts a dimeric solution-state structure, unlike any other quaternary association observed for the DAH7PSs characterised to date. Surprisingly, *Pae*DAH7PS^{PA1901} contains a novel major interface that has not previously been characterised in any DAH7PS. The formation of this alternative major interface in *Pae*DAH7PS^{PA1901}, relative to either of the oligomeric interfaces observed in *Pae*DAH7PS^{PA2843} or *Mtu*DAH7PS, disrupts completely the formation of any aromatic amino acid allosteric binding sites that are comparable with those observed in *Pae*DAH7PS^{PA2843} or *Mtu*DAH7PS. The subsequent insensitivity of *Pae*DAH7PS^{PA1901} to allosteric inhibition by aromatic amino acids is compatible with delivering chorismate to support secondary metabolism, in contrast with *Pae*DAH7PS^{PA2843} or *Mtu*DAH7PS, which are sensitive to either Trp or combinations of aromatic amino acids that include Trp, and function primarily within primary metabolism.

Clear sequence diversity exists between the two type II DAH7PS groups identified by sequence clustering analysis. These different sequence characteristics translate directly into two groups of type II DAH7PSs that form significantly different oligomeric interfaces and quaternary assemblies with associated distinct allosteric functionalities. In addition, these differences in quaternary assembly and allosteric behaviour between the two type II DAH7PS groups relate to their defined physiological roles within either primary or secondary metabolism. On this basis, we propose that there is sufficient diversity between these two groups of type II DAH7PSs, both in terms of primary structure and functionality of the resultant enzymes, that the type II DAH7PSs be further categorised as type II_A and type II_B. The type II_A DAH7PSs comprise full-length enzymes containing both an N-terminal extension and the α_{2a} and α_{2b} helices (for example *Pae*DAH7PS^{PA2843}, *Mtu*DAH7PS or *Cgl*DAH7PS). Type II_A DAH7PS function primarily within primary metabolism, whereas the type II_B DAH7PSs comprise short-form enzymes that contain the N-terminal extension but omit the α_{2a} and α_{2b} helices and these function primarily within secondary metabolism (for example *Pae*DAH7PS^{PA1901}).

Acknowledgements

We thank the beamline scientists at the Australian Synchrotron, Victoria, Australia, for carrying out parts of the research on the MX2 and SAXS/WAXS beamlines.

Competing interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

This work was supported by the Maurice Wilkins Centre for Molecular Biodiscovery; the Biomolecular Interaction Centre; and the New Zealand Marsden Fund [grant number UoC 1105].

Author contribution

O.W.S. and E.J.P. designed the experiments. O.W.S. performed experiments. O.W.S., E.J.M.L., S.A.K., T.M.R., B.D., G.B.J. and E.J.P. analysed the data. O.W.S., G.B.J. and E.J.P. wrote the paper. O.W.S., E.J.M.L., S.A.K., T.M.R., B.D., G.B.J. and E.J.P. edited the paper.

Abbreviations

AUC, analytical ultracentrifugation; *Cgl*, *Corynebacterium glutamicum*; DAH7P, 3-deoxy-D-arabino-heptulosonate 7-phosphate; DAH7PS, DAH7P synthase; DMGA, Discrete Model Genetic Algorithm; E4P, erythrose 4-phosphate; *Mtu*, *Mycobacterium tuberculosis*; *Pae*, *Pseudomonas aeruginosa*; PCA, phenazine-1-carboxylic acid; PDB, Protein Data Bank; PEP, phosphoenolpyruvate; Phe, phenylalanine; PYO, pyocyanin; SAXS, small angle X-ray scattering; SEC-SAXS, size-exclusion chromatography coupled SAXS; Tyr, tyrosine; Trp, tryptophan; TEV, tobacco etch virus protease.

References

- Cohen, T.S. and Prince, A. (2012) Cystic fibrosis: a mucosal immunodeficiency syndrome. *Nat. Med.* **18**, 509–519, <https://doi.org/10.1038/nm.2715>
- Pierson, L.S. and Pierson, E.A. (2010) Metabolism and function of phenazines in bacteria: impacts on the behavior of bacteria in the environment and biotechnological processes. *Appl. Microbiol. Biotechnol.* **86**, 1659–1670, <https://doi.org/10.1007/s00253-010-2509-3>
- Dietrich, L.E.P., Price-Whelan, A., Petersen, A., Whiteley, M. and Newman, D.K. (2006) The phenazine pyocyanin is a terminal signalling factor in the quorum sensing network of *Pseudomonas aeruginosa*. *Mol. Microbiol.* **61**, 1308–1321, <https://doi.org/10.1111/j.1365-2958.2006.05306.x>
- Price-Whelan, A., Dietrich, L.E.P. and Newman, D.K. (2007) Pyocyanin alters redox homeostasis and carbon flux through central metabolic pathways in *Pseudomonas aeruginosa* PA14. *J. Bacteriol.* **189**, 6372–6381, <https://doi.org/10.1128/JB.00505-07>
- Rada, B., Lekstrom, K., Damian, S., Dupuy, C. and Leto, T.L. (2008) The *Pseudomonas* toxin pyocyanin inhibits the dual oxidase-based antimicrobial system as it imposes oxidative stress on airway epithelial cells. *J. Immunol.* **181**, 4883–4893, <https://doi.org/10.4049/jimmunol.181.7.4883>
- Rada, B. and Leto, T.L. (2013) Pyocyanin effects on respiratory epithelium: relevance in *Pseudomonas aeruginosa* airway infections. *Trends Microbiol.* **21**, 73–81, <https://doi.org/10.1016/j.tim.2012.10.004>
- Chang, S., Xu, X., Balsiger, R., Boyaka, P. and Cormet-Boyaka, E. (2015) The pseudomonas aeruginosa toxin pyocyanin decreases expression of cfr and impairs rescue of delta f508-cfr via activation of the mek/erk mapk pathway. *Pediatr. Pulmonol.* **50**, 198
- Lau, G.W., Ran, H.M., Kong, F.S., Hassett, D.J. and Mavrodi, D. (2004) *Pseudomonas aeruginosa* pyocyanin is critical for lung infection in mice. *Infect. Immun.* **72**, 4275–4278, <https://doi.org/10.1128/IAI.72.7.4275-4278.2004>
- Mavrodi, D.V., Ksenzenko, V.N., Bonsall, R.F., Cook, R.J., Boronin, A.M. and Thomashow, L.S. (1998) A seven-gene locus for synthesis is of phenazine-1-carboxylic acid by *Pseudomonas fluorescens* 2-79. *J. Bacteriol.* **180**, 2541–2548
- Schuster, M., Lostroh, C.P., Ogi, T. and Greenberg, E.P. (2003) Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: a transcriptome analysis. *J. Bacteriol.* **185**, 2066–2079, <https://doi.org/10.1128/JB.185.7.2066-2079.2003>
- Deziel, E., Gopalan, S., Tampakaki, A.P., Lepine, F., Padfield, K.E., Saucier, M. et al. (2005) The contribution of Mvfr to *Pseudomonas aeruginosa* pathogenesis and quorum sensing circuitry regulation: multiple quorum sensing-regulated genes are modulated without affecting lasRI, rhlRI or the production of N-acyl-L-homoserine lactones. *Mol. Microbiol.* **55**, 998–1014, <https://doi.org/10.1111/j.1365-2958.2004.04448.x>
- Mavrodi, D.V., Bonsall, R.F., Delaney, S.M., Soule, M.J., Phillips, G. and Thomashow, L.S. (2001) Functional analysis of genes for biosynthesis of pyocyanin and phenazine-1-carboxamide from *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **183**, 6454–6465, <https://doi.org/10.1128/JB.183.21.6454-6465.2001>
- Parsons, J.F., Calabrese, K., Eisenstein, E. and Ladner, J.E. (2003) Structure and mechanism of *Pseudomonas aeruginosa* PhzD, an isochorismatase from the phenazine biosynthetic pathway. *Biochemistry* **42**, 5684–5693, <https://doi.org/10.1021/bi027385d10.1021/bi027385d>
- Parsons, J.F., Song, F.H., Parsons, L., Calabrese, K., Eisenstein, E. and Ladner, J.E. (2004) Structure and function of the phenazine biosynthesis protein PhzF from *Pseudomonas fluorescens* 2-79. *Biochemistry* **43**, 12427–12435, <https://doi.org/10.1021/bi049059z10.1021/bi049059z>
- Parsons, J.F., Calabrese, K., Eisenstein, E. and Ladner, J.E. (2004) Structure of the phenazine biosynthesis enzyme PhzG. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2110–2113, <https://doi.org/10.1107/S0907444904022474>
- Ahuja, E.G., Janning, P., Mentel, M., Graebisch, A., Breinbauer, R., Hiller, W. et al. (2008) PhzA/B catalyzes the formation of the tricyclic in phenazine biosynthesis. *J. Am. Chem. Soc.* **130**, 17053–17061, <https://doi.org/10.1021/ja806325k10.1021/ja806325k>
- Li, Q.A., Mavrodi, D.V., Thomashow, L.S., Roessle, M. and Blankenfeldt, W. (2011) Ligand binding induces an ammonia channel in 2-Amino-2-desoxyisochorismate (ADIC) synthase PhzE. *J. Biol. Chem.* **286**, 18213–18221, <https://doi.org/10.1074/jbc.M110.183418>
- Parsons, J.F., Greenhagen, B.T., Shi, K., Calabrese, K., Robinson, H. and Ladner, J.E. (2007) Structural and functional analysis of the pyocyanin biosynthetic protein PhzM from *Pseudomonas aeruginosa*. *Biochemistry* **46**, 1821–1828, <https://doi.org/10.1021/bi602440310.1021/bi6024403>
- Greenhagen, B.T., Shi, K., Robinson, H., Gamage, S., Bera, A.K., Ladner, J.E. et al. (2008) Crystal structure of the pyocyanin biosynthetic protein PhzS. *Biochemistry* **47**, 5281–5289, <https://doi.org/10.1021/bi702480t10.1021/bi702480t>
- 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, <https://doi.org/10.1099/13500872-142-8-1973>
- Subramaniam, P.S., Xie, G., Xia, T.H. 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

- 22 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*. *Struct. Fold. Des.* **7**, 865–875, [https://doi.org/10.1016/S0969-2126\(99\)80109-9](https://doi.org/10.1016/S0969-2126(99)80109-9)
- 23 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, <https://doi.org/10.1073/pnas.0337566100>
- 24 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, <https://doi.org/10.1016/j.jmb.2004.05.077>
- 25 Schofield, L.R., Anderson, B.F., Patchett, M.L., Norris, G.E., Jameson, G.B. and Parker, E.J. (2005) Substrate ambiguity and crystal structure of *Pyrococcus furiosus* 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase: an ancestral 3-deoxyald-2-ulosonate-phosphate synthase? *Biochemistry* **44**, 11950–11962, <https://doi.org/10.1021/bi050577z>
- 26 Webby, C.J., Baker, H.M., Lott, J.S., Baker, E.N. and Parker, E.J. (2005) The structure of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase from *Mycobacterium tuberculosis* reveals a common catalytic scaffold and ancestry for type I and type II enzymes. *J. Mol. Biol.* **354**, 927–939, <https://doi.org/10.1016/j.jmb.2005.09.093>
- 27 Zhou, L., Wu, J., Janakiraman, V., Shumilin, I.A., Bauerle, R., Kretsinger, R.H. et al. (2012) Structure and characterization of the 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase from *Aeropyrum pernix*. *Bioorg. Chem.* **40**, 79–86, <https://doi.org/10.1016/j.bioorg.2011.09.002>
- 28 Light, S.H., Halavaty, A.S., Minasov, G., Shuvalova, L. and Anderson, W.F. (2012) Structural analysis of a 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase with an N-terminal chorismate mutase-like regulatory domain. *Protein Sci.* **21**, 887–895, <https://doi.org/10.1002/pro.2075>
- 29 Chaudhury, S., Abdulhameed, M.D.M., Singh, N., Tawa, G.J., D'Haeseleer, P.M., Zemla, A.T. et al. (2013) Rapid countermeasure discovery against *Francisella tularensis* based on a metabolic network reconstruction. *PLoS ONE* **8**, 13, <https://doi.org/10.1371/journal.pone.0063369>
- 30 Cross, P.J., Pietersma, A.L., Allison, T.M., Wilson-Coutts, S.M., Cochrane, F.C. and Parker, E.J. (2013) *Neisseria meningitidis* expresses a single 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase that is inhibited primarily by phenylalanine. *Protein Sci.* **22**, 1087–1099, <https://doi.org/10.1002/pro.2293>
- 31 Nazmi, A.R., Lang, E.J.M., Bai, Y., Allison, T.M., Othman, M.H., Panjikar, S. et al. (2016) Interdomain conformational changes provide allosteric regulation en route to chorismate. *J. Biol. Chem.* **291**, 21836–21847, <https://doi.org/10.1074/jbc.M116.741637>
- 32 Burschowsky, D.T., Thorbjørnsrud, H.V., Heim, J.B., Fahrig-Kamarauskaitė, J., Würth-Roderer, K., Kast, P. et al. (2018) Inter-enzyme allosteric regulation of chorismate mutase in *Corynebacterium glutamicum*: structural basis of feedback activation by Trp. *Biochemistry* **57**, 557–573, <https://doi.org/10.1021/acs.biochem.7b01018>
- 33 Sterritt, O.W., Kessans, S.A., Jameson, G.B. and Parker, E.J. (2018) A pseudoisosteric type II DAH7PS enzyme from *Pseudomonas aeruginosa*: alternative evolutionary strategies to control shikimate pathway flux. *Biochemistry* **57**, 78, <https://doi.org/10.1021/acs.biochem.8b00082>
- 34 Webby, C.J., Jiao, W.T., Hutton, R.D., Blackmore, N.J., Baker, H.M., Baker, E.N. et al. (2010) Synergistic allostery, a sophisticated regulatory network for the control of aromatic amino acid biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **285**, 30567–30576, <https://doi.org/10.1074/jbc.M110.111856>
- 35 Jiao, W.T., Hutton, R.D., Cross, P.J., Jameson, G.B. and Parker, E.J. (2012) Dynamic cross-talk among remote binding sites: the molecular basis for unusual synergistic allostery. *J. Mol. Biol.* **415**, 716–726, <https://doi.org/10.1016/j.jmb.2011.11.037>
- 36 Blackmore, N.J., Reichau, S., Jiao, W.T., Hutton, R.D., Baker, E.N., Jameson, G.B. et al. (2013) Three sites and you are out: ternary synergistic allostery controls aromatic amino acid biosynthesis in *Mycobacterium tuberculosis*. *J. Mol. Biol.* **425**, 1582–1592, <https://doi.org/10.1016/j.jmb.2012.12.019>
- 37 Sasso, S., Okvist, M., Roderer, K., Gamper, M., Codoni, G., Kregel, U. et al. (2009) Structure and function of a complex between chorismate mutase and DAHP synthase: efficiency boost for the junior partner. *EMBO J.* **28**, 2128–2142, <https://doi.org/10.1038/emboj.2009.165>
- 38 Blackmore, N.J., Nazmi, A.R., Hutton, R.D., Webby, M.N., Baker, E.N., Jameson, G.B. et al. (2015) Complex formation between two biosynthetic enzymes modifies the allosteric regulatory properties of both an example of molecular symbiosis. *J. Biol. Chem.* **290**, 18187–18198, <https://doi.org/10.1074/jbc.M115.638700>
- 39 Webby, C.J., Patchett, M.L. and Parker, E.J. (2005) Characterization of a recombinant type II 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Helicobacter pylori*. *Biochem J.* **390**, 223–230, <https://doi.org/10.1042/BJ20050259>
- 40 Finn, R.D., Bateman, A., Clements, J., Coghill, P., Eberhardt, R.Y., Eddy, S.R. et al. (2014) Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222–D30, <https://doi.org/10.1093/nar/gkt1223>
- 41 Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W.Z. et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 6
- 42 Finn, R.D., Clements, J. and Eddy, S.R. (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29–W37, <https://doi.org/10.1093/nar/gkr367>
- 43 Apweiler, R., Bateman, A., Martin, M.J., O'Donovan, C., Magrane, M., Alam-Faruque, Y. et al. (2014) Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **42**, D191–D198, <https://doi.org/10.1093/nar/gkt1140>
- 44 Sayers, E.W., Barrett, T., Benson, D.A., Bolton, E., Bryant, S.H., Canese, K. et al. (2011) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **39**, D38–D51, <https://doi.org/10.1093/nar/gkq1172>

- 45 Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. and Barton, G.J. (2009) Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191, <https://doi.org/10.1093/bioinformatics/btp033>
- 46 Frickey, T. and Lupas, A. (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* **20**, 3702–3704, <https://doi.org/10.1093/bioinformatics/bth444>
- 47 Demeler, B. and van Holde, K.E. (2004) Sedimentation velocity analysis of highly heterogeneous systems. *Anal. Biochem.* **335**, 279–288, <https://doi.org/10.1016/j.ab.2004.08.039>
- 48 Brookes, E.H. and Demeler, B. (2007) Parsimonious regularization using genetic algorithms applied to the analysis of analytical ultracentrifugation experiments. *Gecco 2007: Genetic and Evolutionary Computation Conference* **1–2**, 361–368
- 49 Demeler, B. and Brookes, E. (2008) Monte Carlo analysis of sedimentation experiments. *Colloid Polym. Sci.* **286**, 129–137, <https://doi.org/10.1007/s00396-007-1699-4>
- 50 Demeler, B. (2010) Methods for the design and analysis of sedimentation velocity and sedimentation equilibrium experiments with proteins. *Curr. Protoc. Protein Sci.*, <https://doi.org/10.1002/0471140864.ps0713s60>
- 51 Rocco, M. and Byron, O. (2015) Computing translational diffusion and sedimentation coefficients: an evaluation of experimental data and programs. *Eur. Biophys. J. Biophys. Lett.* **44**, 417–431, <https://doi.org/10.1007/s00249-015-1042-9>
- 52 Brookes, E., Vachette, P., Rocco, M. and Perez, J. (2016) US-SOMO HPLC-SAXS module: dealing with capillary fouling and extraction of pure component patterns from poorly resolved SEC-SAXS data. *J. Appl. Crystallogr.* **49**, 1827–1841, <https://doi.org/10.1107/S1600576716011201>
- 53 Kirby, N.M., Mudie, S.T., Hawley, A.M., Cookson, D.J., Mertens, H.D.T., Cowieson, N. et al. (2013) A low-background-intensity focusing small-angle X-ray scattering undulator beamline. *J. Appl. Crystallogr.* **46**, 1670–1680, <https://doi.org/10.1107/S002188981302774X>
- 54 Kirby, N., Cowieson, N., Hawley, A.M., Mudie, S.T., McGillivray, D.J., Kusel, M. et al. (2016) Improved radiation dose efficiency in solution SAXS using a sheath flow sample environment. *Acta Crystallogr. Sect. D Struct. Biol.* **72**, 1254–1266, <https://doi.org/10.1107/S2059798316017174>
- 55 Ryan, T.M., Trehwella, J., Murphy, J.M., Keown, J.R., Casey, L., Pearce, F.G. et al. (2018) An optimized SEC-SAXS system enabling high X-ray dose for rapid SAXS assessment with correlated UV measurements for biomolecular structure analysis. *J. Appl. Crystallogr.* **51**, 97–111, <https://doi.org/10.1107/S1600576717017101>
- 56 Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. and Svergun, D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282, <https://doi.org/10.1107/S0021889803012779>
- 57 Brookes, E., Perez, J., Cardinali, B., Profumo, A., Vachette, P. and Rocco, M. (2013) Fibrinogen species as resolved by HPLC-SAXS data processing within the UltraScan Solution Modeller (US-SOMO) enhanced SAS module. *J. Appl. Crystallogr.* **46**, 1823–1833, <https://doi.org/10.1107/S0021889813027751>
- 58 Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M. et al. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350, <https://doi.org/10.1107/S0021889812007662>
- 59 Chayen, N.E. (1997) A novel technique to control the rate of vapour diffusion, giving larger protein crystals. *J. Appl. Crystallogr.* **30**, 198–202, <https://doi.org/10.1107/S0021889896013532>
- 60 McPhillips, T.M., McPhillips, S.E., Chiu, H.J., Cohen, A.E., Deacon, A.M., Ellis, P.J. et al. (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron. Radiat.* **9**, 401–406, <https://doi.org/10.1107/S0909049502015170>
- 61 Kabsch, W. (2010) XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132, <https://doi.org/10.1107/S0907444909047337>
- 62 Evans, P.R. and Murshudov, G.N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214, <https://doi.org/10.1107/S0907444913000061>
- 63 Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022–1025, <https://doi.org/10.1107/S0021889897006766>
- 64 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132, <https://doi.org/10.1107/S0907444904019158>
- 65 Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53**, 240–255, <https://doi.org/10.1107/S0907444996012255>
- 66 Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797, <https://doi.org/10.1016/j.jmb.2007.05.022>
- 67 Kabsch, W. (1976) Solution for best rotation to relate 2 sets of vectors. *Acta Crystallogr. Sect. A* **32**, 922–923, <https://doi.org/10.1107/S0567739476001873>
- 68 Liu, Y.J., Li, P.P., Zhao, K.X., Wang, B.J., Jiang, C.Y., Drake, H.L. et al. (2008) *Corynebacterium glutamicum* contains 3-deoxy-D-arabino-heptulosonate 7-phosphate synthases that display novel biochemical features. *Appl. Environ. Microbiol.* **74**, 5497–5503, <https://doi.org/10.1128/AEM.00262-08>
- 69 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 Expr. Purif.* **34**, 17–27, <https://doi.org/10.1016/j.pep.2003.11.008>
- 70 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, [https://doi.org/10.1016/S0022-2836\(02\)00545-4](https://doi.org/10.1016/S0022-2836(02)00545-4)

- 71 Jiao, W., Blackmore, N.J., Nazmi, A.R. and Parker, E.J. (2017) Quaternary structure is an essential component that contributes to the sophisticated allosteric regulation mechanism in a key enzyme from *Mycobacterium tuberculosis*. *PLoS ONE* **12**, 21, <https://doi.org/10.1371/journal.pone.0180052>
- 72 Fischer, H., Neto, M.D., Napolitano, H.B., Polikarpov, I. and Craievich, A.F. (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J. Appl. Crystallogr.* **43**, 101–109, <https://doi.org/10.1107/S0021889809043076>

Structural and Functional Characterisation of the
Entry Point to Pyocyanin Biosynthesis in
Pseudomonas aeruginosa Defines a New 3-
Deoxy-D-arabino-heptulosonate 7-phosphate
Synthase Subclass

*Oliver W. Sterritt^{1,2}, Eric J. M. Lang^{1,#}, Sarah A. Kessans¹, Timothy M. Ryan³, Borries
Demeler⁴, Geoffrey B. Jameson^{1,2,5,*} and Emily J. Parker^{1,2,6,*}*

¹Biomolecular Interaction Centre and School of Physical and Chemical Sciences, University of Canterbury, Christchurch 8041, New Zealand

²Maurice Wilkins Centre for Molecular Biodiscovery, University of Auckland 1142, New Zealand

³SAXS/WAXS, Australian Synchrotron/ANSTO, 800 Blackburn Road, Clayton, VIC 3168, Australia

⁴Department of Chemistry and Biochemistry, University of Lethbridge, Alberta T1K 3M4, Canada

⁵Institute of Fundamental Sciences and the Riddet Institute, Massey University, Palmerston North 4442, New Zealand

⁶Ferrier Research Institute, Victoria University of Wellington, Wellington 6012, New Zealand

[#]Present address: School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom

SUPPORTING INFORMATION

MtuDAH7PS β_0 α_{0a} α_{0b}

1 10 20 30 40 50

MtuDAH7PS MNWTVVDIPIDQLPSLPLPTD L R L R T I D A L A K P A A O O P T W P A D Q A M A M R T V L E S V P P V T
PaeDAH7PS_PA2843 M S Q S W S P E S W R A K P I Q O O P E Y P D A A E H T A R V E Q T L A G Y P P L V
PaeDAH7PS_PA1901 M D D L L Q R V R R C E A I O O P E W G D P S R T R D V Q A Y T R G S P A L I
SacDAH7PS M S R Y P F P G A G H D P L I A S M V A A L P A R O O P Q W P E P S R T A P V H E R L A R E S F L V
PchDAH7PS M E D I L K R V L S C E A F O O P Q W S E P S Q T H D A Q A Y L R D S A S L I
SfuDAH7PS_SFUL3326 M D N V V S E I R L R P A I O O P D W P D A E R T E E V R R T L S G R R A L V
SfuDAH7PS_SFUL5264 M D I D V R D I S S R V A L O O P D W A D D P R T E V R A Q T A R H D A L V

MtuDAH7PS α_{0c} β_1 α_1

60 70 80 90 100 110

MtuDAH7PS V P S E I V R L Q E Q L A Q V A K G E A F L L Q G G D C A E T F M D N T E P H I R G N V R A L L Q M A V V L T Y G A S M
PaeDAH7PS_PA2843 F A G E A R E L R R Q F A E V T A G R A F L L Q G G D C A E S F A E F S A A K I R D T F R V I L Q M A V V M T F A A G C
PaeDAH7PS_PA1901 R A C D I L A L R A T L A R V A R G E A L V V Q G G C A E D M D H H A E N V A R K A V I E L L A G A L R L A G R R
SacDAH7PS S Y T S V R E L R G L L A R A A E C E F C L I Q A G D C V E L T T E C E P C D V E R K V E M D D V L G D V M R T G S G L
PchDAH7PS R Y E D I L V L R A T L A R V A A C E A M V I Q S G D C A E D M D E S T P D H V V R K A V I L L A G T F R I V T Q Q
SfuDAH7PS_SFUL3326 R Y E D V R D L R Q I L A R V A R G E A Q V V Q G G C A E D P M E S S A G Y V A R K A A V I D L L A G A M K M A S H K
SfuDAH7PS_SFUL5264 R A E D V R R L R A V L A R V A A G E A H L V L A G D C A E D P A E C T S S Y V A R K A A I D V L A G T L K M V T R R

MtuDAH7PS β_2 TT

120 130 140 150 160 170

MtuDAH7PS P V V K V A R I A G O Y A K P R S A D I D A L G L R S Y R G D M I N G F A P D A A A R E H D P S R L V R A Y A N A
PaeDAH7PS_PA2843 P V V K V G R M A G O F A K P R S S G D E T Q N G V T L P A Y R G D I V N G I G F D E K S R V P D P E R L L Q A Y H Q S
PaeDAH7PS_PA1901 P V I R V G R I A G O F A K P R S K P H E Q V G E Q T L P V Y R G D M V N G R E A H A E Q R A D P P O R I L G Y A A A
SacDAH7PS P V V R V G R M A G O F A K P R S D D W E T V E G G R L P V F R G P V V N A P D A C E E A R T F P S R V I T G L E A A
PchDAH7PS P V V R V G R I A G O F A K P R S N N N E R I G D V E L P V Y R G D M V N G R E A I C G R Q H D A Q R L V R G Y S A A
SfuDAH7PS_SFUL3326 P I V R I G R I A G O F A K P R S Q P V E Y V G G V E L P V Y R G H M V N S P E P A P K G R I P D P S R L L T C Y D A A
SfuDAH7PS_SFUL5264 P V V R V G R M A G O F A K P R S S P T E R V C G Q E L P V F R G H M V N G P E P D P V E R A D P T R I L S C Y R A A

MtuDAH7PS α_2 α_{2a} α_{2b}

180 190 200 210 220 230

MtuDAH7PS S A A M N I V R A L T S S G L A S I H L V H D W N R E F V R T S P A G A R Y E A L A T E I D R G I R F M S A C G V A D R
PaeDAH7PS_PA2843 T A S L N L L R A F A Q G G F A D L H Q V H R W N L D F I A N S A L A E R Y Q O L A D R I D E T L A F M R A C G L D S A
PaeDAH7PS_PA1901 R N I M R H L G W D A A S G
SacDAH7PS R C A V T A L D R L G R G E
PchDAH7PS R D I M Q H L G W K A S A S Q
SfuDAH7PS_SFUL3326 G E M M G H L G W H S T A R E F
SfuDAH7PS_SFUL5264 A A V M A G L G R L P S G A

MtuDAH7PS β_3 α_3 β_4 α_4

240 250 260 270 280 290

MtuDAH7PS . N L Q T A E I Y A S H E A L V L D Y E R A M L S L S D G D D G E P Q L F D L S A H T V W I G E R T R O I D C A H T A F
PaeDAH7PS_PA2843 P Q L R E T S F F T A H E A L L N Y E A L T R R D S I T C E W Y D C S A H M L W I G D R T R O I D C A H V E M
PaeDAH7PS_PA1901 E . A N A S P V W T S H E M L L D Y E L S M R E D E Q R R V Y L G S T H W P W I G E R T R O V D C A H V A L
SacDAH7PS G A A P E E R V W T S H E A L L D Y E L P Q V R R H T D C G S Y I A S T H W P W I G E R T R O A D A H V R L
PchDAH7PS E Q L S G S P A W T S H E M L V L D Y E L P Q R R D E Q C R V F L G S T H W P W I G E R T R O L T C A H V T L
SfuDAH7PS_SFUL3326 G R A I D P P V W T S H E A L L D Y E L P M L R R D E S G A R W L A S T H L P W I G E R T R A L D C A H V E L
SfuDAH7PS_SFUL5264 P R Q I D P D V W T S H E A L L D Y E V P M L R R E A G C Q L L I G S T H W P W I G E R T R O V E C A H V A L

MtuDAH7PS β_5 α_5 β_6 α_6

300 310 320 330 340 350

MtuDAH7PS A Q V I A N P V G V K L G P N M T P E L A V E Y V E R L D P H N K P G R I T L V S R M G N H K V R D L L P P I V E K V Q
PaeDAH7PS_PA2843 I R G V G N P I G V K V G S M D S E E L I R L I D I L N P D N D P G R I N L I V R M G A D K V G D H I P R L I Q A I Q
PaeDAH7PS_PA1901 L A E V L N P V A C K V G P E I G R D Q L L A L C E R L D P R R E P G R I T L I A R M G A Q V G E R L P P I V E A V R
SacDAH7PS M T E L D N P V A C K V G P E A T V D E V L D L C A V L D P H R T P G R I T L I A R F G A D R V A . A L A P L V R A V R
PchDAH7PS I S E V I N P V A C K V G P D I T R D Q L L S L C E R L D S K R E P G R I T L I A R M G A Q V A D R L P P I V E A V R
SfuDAH7PS_SFUL3326 F A S I A N P V A C K V G P S M T V A D L L G L C E R L D P D R E P G R I T L I S R M G A G A V G E R L P A L V A A V R
SfuDAH7PS_SFUL5264 L A A V V N P V A C K V G P R M A P D E L V A L C A R L D P G R E P G R I T L I A R M G A G A V A G A L P P I V T A V R

MtuDAH7PS β_7 α_7 β_8

360 370 380 390 400 410

MtuDAH7PS A T C H Q V I W Q C D P M H G N T H E S S T G K T R H F D R I V D E V Q G F F E V H R A L G T H P G G I H V E I T G E
PaeDAH7PS_PA2843 R E C R Q V I W S S D P M H G N T I K A S S G Y K T R D F A R V L A E V R Q F F E V H Q A E G S Y A G G I H E I M T G Q
PaeDAH7PS_PA1901 A A C H P V I W L S D P M H G N T I V A P C G N K T R L V R S I A E V A F R L A V S G S G V A A G L H E I T T P D
SacDAH7PS Q A C H P V L W M C D P M H G N T V K T P D G L K T R R L D T I M S E I R Q C V D V L A E N G E R C A G L H E I A S P D
PchDAH7PS Q A C H K I I W L S D P M H G N T I V A P C G N K T R M V Q A I T E I T A F K H A V T S A G G V A A G L H E I T T P D
SfuDAH7PS_SFUL3326 A A C H P V S W L C D P M H G N T V T T G E G L K T R Y L E H V E R E V R G F L T A V R S A D G T A G G I H E I T T P E
SfuDAH7PS_SFUL5264 E A C H P V I W L T D P M H G N T V T T A T G V K T R F V E T V V E E V R Q F Q L A V R Q A G G V A G L H E I A T P D

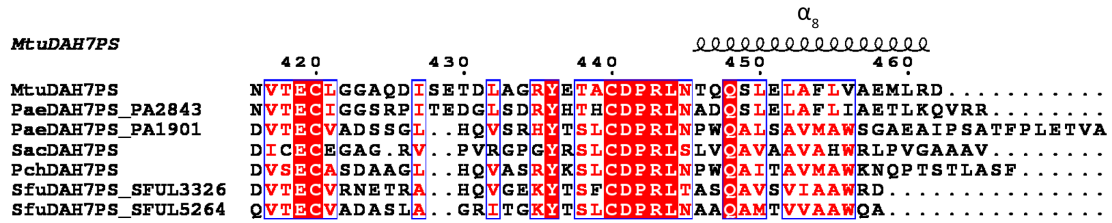


Figure S1: Sequence Alignment among the type II DAH7PSs, indicating the relative positions of β strands and α helices based on *Mtu*DAH7PS (PDB 2B7O). *Mtu*, *Mycobacterium tuberculosis*; *Pae_PA2843*, *Pseudomonas aeruginosa*; *Pae_PA1901*, *Pseudomonas aeruginosa*; *Sac*, *Streptomyces achromogenes*; *Pch*, *Pseudomonas chlororaphis*; *Sfu_SFUL3326*, *Streptomyces fulvissimus*; *Sfu_SFUL5264*, *Streptomyces fulvissimus*.

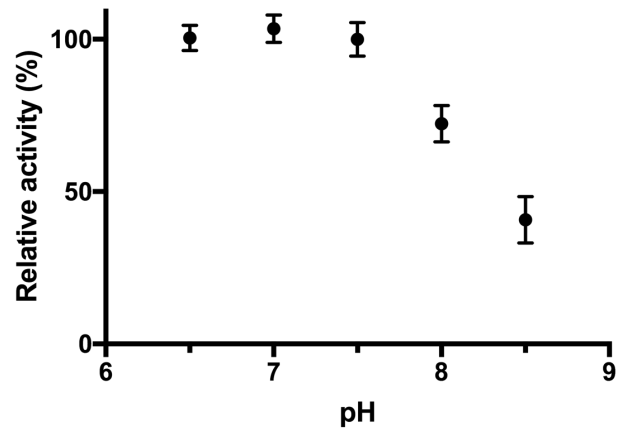


Figure S2: Activity of *PaeDAH7PS*^{PA1901} measured at a range of pH, relative to the activity observed at pH 7.5, indicates that this enzyme is active over a range of pH between pH 6.5 and 7.5. Substrate concentrations were held at 150 μ M each for PEP and E4P.

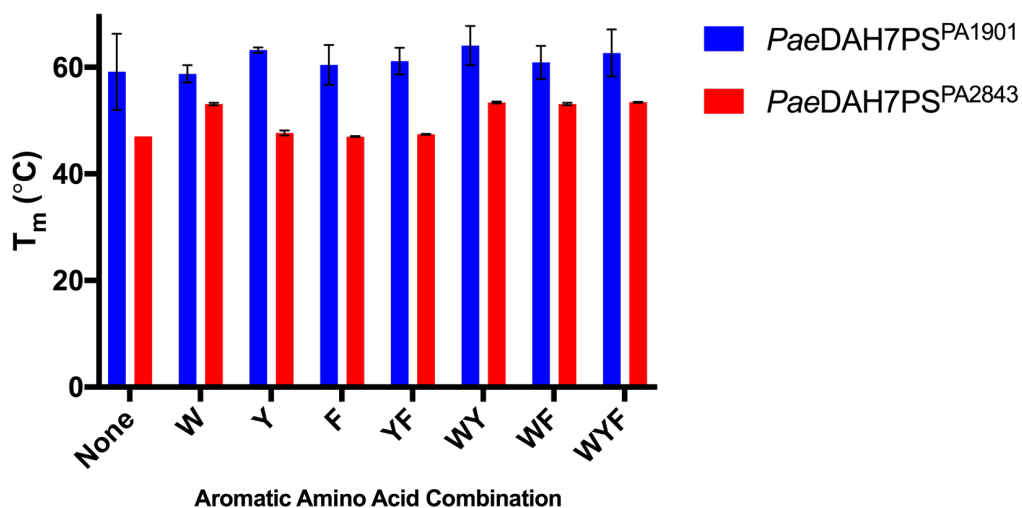


Figure S3: The melting temperature of *PaeDAH7PS*^{PA1901} and comparison to that previously reported for *PaeDAH7PS*^{PA2843} [1] in the presence of various combinations of aromatic amino acids. Each single letter code corresponds to 200 μ M of the appropriate amino acid.

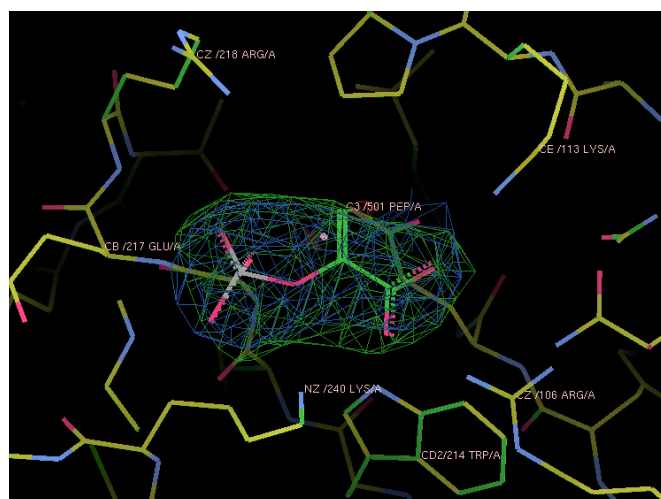


Figure S4: Electron density omit maps for PEP at the active site of *PaeDAH7PS*^{PA1901} (PDB **6BMC**). 2IFol-IFcl is represented by the blue mesh contoured at 1.5 σ and IFol-IFcl is represented by the green mesh contoured at 3 σ .

Table S1: SEC-SAXS sample and data collection parameters.

Sample Details	<i>PaeDAH7PS</i>^{PA1901}
Organism	<i>Pseudomonas aeruginosa</i>
UniProt sequence ID	G3XCJ9
Extinction coefficient [A_{280} , 0.1 %]	1.22
M from chemical composition (Da)	44, 468
SEC-SAXS column	5 x 150 mm, Superdex S200
Loading concentration (mg mL ⁻¹)	1.0, 5.0 or 8.0
Injection volume (μ L)	20
Flow rate (mL min ⁻¹)	0.2
Solvent	50 mM bis-Tris propane pH 7.5, 100 μ M cobalt chloride, 200 μ M PEP, 5 % glycerol
SAXS data collection parameters	
Instrument	Australian synchrotron SAXS/WAXS beamline
Detector	Dectris PILATUS 1M
Sample environment	SEC-SAXS with sheath-flow cell
Wavelength (\AA)	1.0332
Camera length (mm)	1600
Set temperature ($^{\circ}$ C)	20

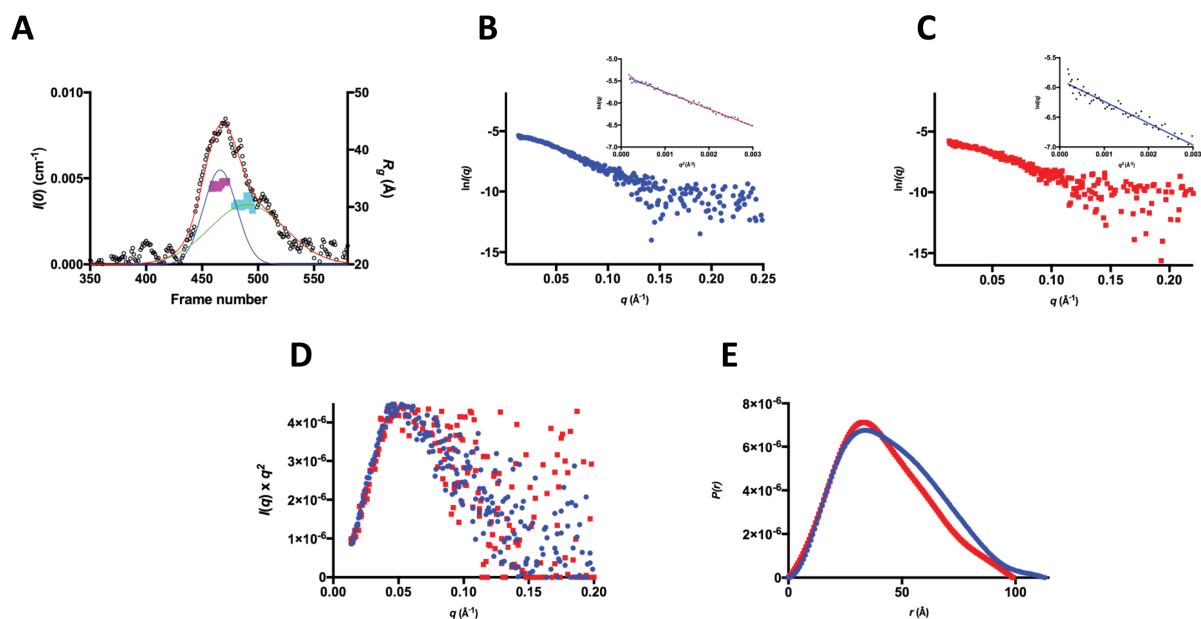


Figure S5: SEC-SAXS analysis for *PaeDAH7PS*^{PA1901} from the 5.0 mg/mL SEC-input concentration. (A) Deconvolution of the SEC-SAXS data indicates two Gaussian components (peak A, blue line and peak B, green line. Sum, red line). The R_g values across each peak are indicated as magenta or cyan squares respectively. (B) – (C) the SAXS profiles for the deconvoluted data peak A and peak B respectively. Guinier plots are inset. (D) Kratky plots of the deconvoluted data in (B) and (C) (peak A, blue circles and peak B, red squares). (E) $P(r)$ plots for the deconvoluted data in (B) and (C) (peak A, blue circles and peak B, red squares).

Table S2: SEC-SAXS parameters determined for *PaeDAH7PS*^{PA1901} from the 5.0 mg/mL SEC-input concentration.

	5.0 mg/mL Peak A	5.0 mg/mL Peak B
<i>Guinier Analysis</i>		
R_g (Å)	34.0 ± 1.6	32.0 ± 3.8
$I(0)$ (cm ⁻¹)	$0.004 \pm 5e^{-5}$	$0.004 \pm 1e^{-4}$
q_{\min}	0.013	0.013
Correlation coefficient, r^2	0.992	0.928
<i>Pair Wise Distribution Analysis</i>		
R_g (Å)	35.25	32.5
$I(0)$ (cm ⁻¹)	0.00	0.00
d_{\max} (Å)	113.1	99.0
V_p (Å ³)	167,000	126,000
q range (Å ⁻¹)	0.013 – 0.25	0.013 – 0.25

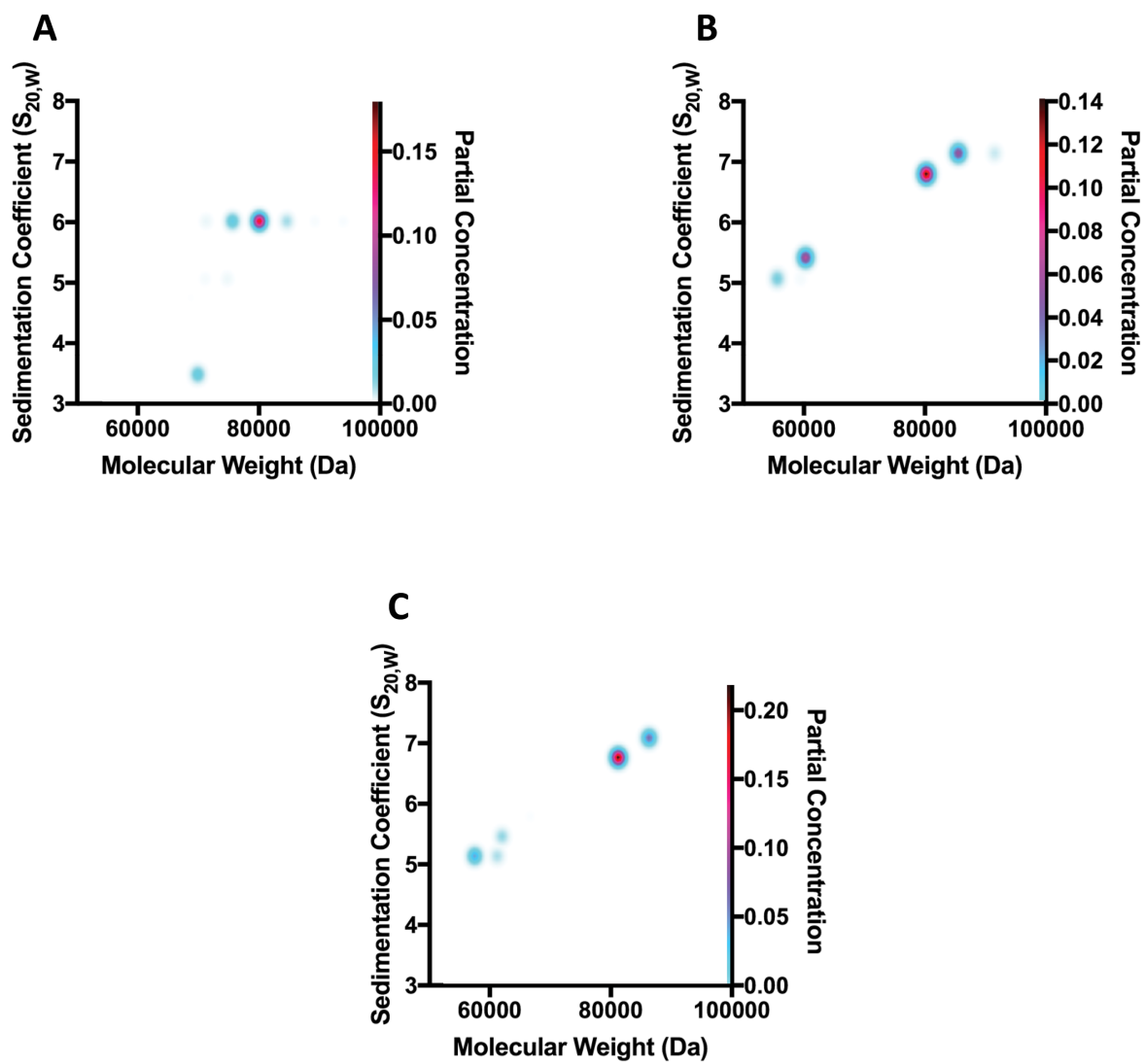


Figure S6: Pseudo-3D plots for distributions of the 2DSA Monte Carlo results shown in Figure 10 for (A) 8 μ M, (B) 23 μ M, and (C) 30 μ M concentrations of *PaeDAH7PS*_{PA1901}.

Table S3: RMSD values obtained for the alignment between the structures of *PaeDAH7PS*^{PA1901} and either *PaeDAH7PS*^{PA2843} or *MtuDAH7PS*.

Alignment between		RMSD (Å)
Single chain		
6BMC_chain A	5UXO	0.816
6BMC_chain A	5UXN	0.830
6BMC_chain A	5UXM	0.822,
6BMC_chain A	2B7O_Chain A	0.775
6BMC_chain A	2B7O_Chain B	0.821
6BMC_chain B	5UXO	0.816
6BMC_Chain B	5UXN	0.817
6BMC_chain B	5UXM	0.840
6BMC_chain B	2B7O_Chain A	0.754
6BMC_chain B	2B7O_Chain B	0.776
Crystallographic Dimer		
6BMC	5UXO	25.740
6BMC	5UXN	26.020
6BMC	5UXM	26.100
6BMC	2B7O	25.100
Crystallographic Tetramer		
6BMC	5UXO	27.850
6BMC	5UXN	28.610
6BMC	5UXM	28.800
6BMC	2B7O	34.460

References

1. Sterritt OW, Kessans, S. A., Jameson, G. B., Parker, E. J. A Pseudoisostructural Type II DAH7PS Enzymes from *Pseudomonas aeruginosa*: alternative evolutionary strategies to control shikimate pathway flux. *Biochemistry*. 2018;57(18):2667 - 78.