Swit_4259, an acetoacetate decarboxylase-like enzyme from *Sphingomonas wittichii* RW1

Lisa S. Mydy,^a Zahra Mashhadi,^b T. William Knight,^a Tyler Fenske,^a Trevor Hagemann,^a Robert W. Hoppe,^a Lanlan Han,^a Todd R. Miller,^c Alan W. Schwabacher^a and Nicholas R. Silvaggi^{a*}

Received 18 September 2017

Accepted 31 October 2017

Edited by I. Tanaka, Hokkaido University, Japan

Keywords: acetoacetate decarboxylase-like enzyme; aldolase; biodegradation; dehydratase; enzyme structure; enzyme mechanism; enzyme catalysis; X-ray crystallography.

PDB reference: Swit_4259, 5upb

Supporting information: this article has supporting information at journals.iucr.org/f

^aDepartment of Chemistry and Biochemistry, University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA, ^bClinical Pharmacology Division, Vanderbilt University, TN 37232, USA, and ^cJoseph J. Zilber School of Public Health, University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA. *Correspondence e-mail: silvaggi@uwm.edu

The Gram-negative bacterium *Sphingomonas wittichii* RW1 is notable for its ability to metabolize a variety of aromatic hydrocarbons. Not surprisingly, the *S. wittichii* genome contains a number of putative aromatic hydrocarbon-degrading gene clusters. One of these includes an enzyme of unknown function, Swit_4259, which belongs to the acetoacetate decarboxylase-like superfamily (ADCSF). Here, it is reported that Swit_4259 is a small (28.8 kDa) tetrameric ADCSF enzyme that, unlike the prototypical members of the superfamily, does not have acetoacetate decarboxylase activity. Structural characterization shows that the tertiary structure of Swit_4259 is nearly identical to that of the true decarboxylases, but there are important differences in the fine structure of the Swit_4259 active site that lead to a divergence in function. In addition, it is shown that while it is a poor substrate, Swit_4259 can catalyze the hydration of 2-oxo-hex-3-enedioate to yield 2-oxo-4-hydroxyhexanedioate. It is also demonstrated that Swit_4259 has pyruvate aldolase-dehydratase activity, a feature that is common to all of the family V ADCSF enzymes studied to date. The enzymatic activity, together with the genomic context, suggests that Swit_4259 may be a hydratase with a role in the metabolism of an as-yet-unknown hydrocarbon. These data have implications for engineering bioremediation pathways to degrade specific pollutants, as well as structure–function relationships within the ADCSF in general.

1. Introduction

Some of the most dangerous and widespread pollutants on earth are aromatic hydrocarbons such as dioxins, polychlorinated biphenyls (PCBs), herbicides and pesticides. These compounds are recalcitrant to degradation in the environment owing to their low solubility in water, their tendency to adsorb to solids and their dissimilarity to naturally occurring compounds (for example halogenation) (Halden & Dwyer, 1997). Thus, there is significant interest in natural or engineered bioremediation pathways that are capable of metabolizing or mineralizing these contaminants to harmless compounds. Some microorganisms, particularly bacteria, have evolved the ability to consume aromatic hydrocarbons as carbon and energy sources (Monna *et al.*, 1993; Aly *et al.*, 2008; Li *et al.*, 2009). *Sphingomonas wittichii* RW1 is a Gram-negative bacterium that was first isolated from contaminated sediments in the Elbe river in Germany, and it is known for its ability to degrade a wide variety of aromatic substrates, including dioxins and PCBs (Wittich *et al.*, 1992; Nam *et al.*, 2014). The *S. wittichii* genome thus contains a large number of aromatic hydrocarbon catabolism gene clusters, of which only a handful have been characterized (Colquhoun *et al.*, 2012; Chai *et al.*, 2016).

Relative to other species, the *S. wittichii* genome is enriched in TonB-dependent receptors (COG1629), dehydrogenases (COG1028) and phenylpropionate dioxygenases (COG4638) (Miller *et al.*, 2010). In fact, the *S. wittichii* genome contains the largest number of TonB-dependent receptors of any genome sequenced to date (Miller *et al.*, 2010; Tang *et al.*, 2012). This is interesting because the TonB system is normally associated with the uptake of iron-bound siderophores, but the majority of the TonB-dependent receptor genes in *S. wittichii* are not in operons that are associated with iron transport. TonB-dependent receptors have been implicated in the transport of other compounds as well, including sugars (Blanvillain *et al.*, 2007), alginate (Postle & Kadner, 2003) and, most importantly for the discussion at hand, aromatic hydrocarbons (Nguyen *et al.*, 2014). This strongly suggests that many of the TonB-dependent receptor-containing gene clusters in *S. wittichii* may actually be cryptic aromatic hydrocarbon-degrading operons.

One of these TonB-dependent receptor-containing, putative aromatic hydrocarbon-degrading gene clusters in *S. wittichii* (Fig. 1) is particularly interesting owing to its inclusion of an acetoacetate decarboxylase-like enzyme, Swit_4259. Despite relatively low amino-acid sequence identity to the prototypical *Clostridium acetobutylicum* acetoacetate decarboxylase (AAB53232; 18%), Swit_4259 is identified by the NCBI Conserved Domain Database (CDD) as a member of the acetoacetate decarboxylase-like superfamily (ADCSF; PF06314). The ADCSF is an incompletely characterized group of enzymes with significant diversity in both substrate and reaction specificities (Burroughs *et al.*, 2013; Mueller *et al.*, 2015). Based on sequence alignments and manual interrogation of putative active-site residues, the ADCSF can be divided into seven families (Burroughs *et al.*, 2013). The characterized acetoacetate decarboxylases (for example the *C. acetobutylicum* enzyme) comprise the majority of ADCSF sequences and are designated family I. The prototypical reaction catalyzed by family I is shown in Supplementary Fig. S1. The reactive lysine residue that forms a Schiff-base intermediate is conserved in all but one ADCSF family (family II). Swit_4259 belongs to ADCSF family V, the members of which, based on gene context, appear to be primarily involved in secondary metabolism. MppR from *Streptomyces hygroscopicus* (also in family V; Magarvey *et al.*, 2006) was shown to catalyze both aldol condensation followed by dehydration of

pyruvate and imidazole-4-carboxaldehyde and the cyclization of 4-hydroxy-2-ketoarginine to produce 2-keto-enduracididine (Burroughs *et al.*, 2013; Supplementary Fig. S1). Another related enzyme, the aldolase-dehydratase from *S. bingchengensis* (SbAD), was shown to be an efficient catalyst of the aldol condensation of pyruvate with a number of unsaturated aliphatic and aromatic aldehydes (Mueller *et al.*, 2015; Supplementary Fig. S1). Here, we describe the tertiary structure of Swit_4259 as well as a preliminary characterization of its enzymatic activity.

2. Materials and methods

2.1. Macromolecule production

The gene at locus Swit_4259 encoding the protein identified by GenBank accession code WP_012050437.1 (Miller *et al.*, 2010) was optimized for expression in *Escherichia coli* and synthesized by GenScript (Piscataway, New Jersey, USA). The Swit_4259 gene was amplified using oligonucleotide primers Swit_4259-pSuf and Swit_4259-pSur. The purified PCR products were digested with BsaI and XbaI and ligated into the pE-SUMO_{kan} expression vector (LifeSensors Inc., Malvern, Pennsylvania, USA) to generate pSUMO-Swit_4259. This plasmid expresses Swit_4259 with a His₆-tagged small ubiquitin-like modifier (SUMO) fused to the N-terminus (Wang *et al.*, 2010). The sequence of pSUMO-Swit_4259 was confirmed by sequencing (GeneWiz Inc.). *E. coli* BL21 Star (DE3) cells (Invitrogen) were transformed with pSUMO-Swit_4259 and used to inoculate 20 ml cultures of Luria–Bertani broth supplemented with 50 µg ml^{−1} kanamycin. These starter cultures were grown overnight at 37°C with agitation at 250 rev min^{−1}. Large-scale cultures (1 l) were inoculated from the starter cultures and grown at 37°C and 225 rev min^{−1} until they reached an OD_{600 nm} of about 1.0. The temperature of the incubator was reduced to 25°C and expression of Swit_4259 was induced by the addition of isopropyl β-D-1-thiogalactopyranoside to a final concentration of 400 µM. Cultures were grown overnight at 25°C. The cells were harvested from the expression culture by centrifugation at 7400g for 30 min and resuspended in buffer A (25 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole) at a ratio of 5 ml buffer per gram of cell pellet [5:1(v:w)]. Lysozyme was added

Figure 1

Gene context of the Swit_4259 locus in *S. wittichii*. The Swit_4259 locus is highlighted in dark blue. Other genes proposed to be co-transcribed with Swit_4259 are colored light blue; genes expected to be outside the operon are colored grey. The annotations below each locus were taken from the GenBank database.

Table 1

Macromolecule-production information.

Source organism	<i>S. wittichii</i> RW1
DNA source	Genomic DNA
Forward primer (Swit_4259-pSUF)†	5′-GGTCTCAAGGTATGGCTGGTGCTACGCA T-3′
Reverse primer (Swit_4259-pSUR)‡	5′-GCTCTAGATCATCACATACGCAGGGATT CCA-3′
Cloning vector	pUC57
Expression vector	pE-SUMO _{kan}
Expression host	<i>E. coli</i> BL21 (DE3)
Complete amino-acid sequence of the construct produced	MAGATHFLPTGQASLVDDALYGWGADMLT VYLRCDFARLQALLPAGLKVADGLCMAYV GAFQSTSEDQPAAMLRNPAGAVYNEAALS IACHTGDRQGYFPAFVWVDKEWSLIRGWL NGYPKKIGAITLARHPYNPVTGGLREGA VVGICARHGFTLFRGLTVTRAGDAGDL RSRPATFGHRHWPALHPTQTPVSELVEVN RSDLRVGDIWAGEPFIELGSAPDEALECF ADHEVLAVGVTYSYGFRI GGATRLSLRM

† The BsaI site is underlined. ‡ The XbaI site is underlined and the stop codon is italicized.

to a final concentration of 1 mg ml⁻¹ and the mixture was stored overnight at -20°C. The cell suspension was thawed at room temperature and allowed to stand for 2 h, after which deoxyribonuclease I (DNase I, Worthington Biochemical Inc.) was added at a ratio of 0.5 mg per 100 ml of lysate. The lysate was clarified by centrifugation at 39 000g for 45 min.

Immobilized metal-affinity chromatography (IMAC) purification of Swit_4259 was performed in two steps. The clarified lysate was applied onto a 5 ml HisTrap column (GE Life Sciences, USA) and eluted using a four-step gradient from 0 to 5, 15, 50 and 100% buffer B (25 mM Tris pH 8.0, 300 mM NaCl, 250 mM imidazole). Swit_4259 eluted primarily at 50% buffer B (about 125 mM imidazole), but a significant amount eluted at the 100% buffer B step (about 250 mM imidazole). The fractions from both elution peaks containing Swit_4259, as judged by Coomassie-stained SDS-PAGE gels, were pooled. SUMO protease (LifeSensors Inc.) was added to a final concentration of about 5 µM (Malakhov *et al.*, 2004). The pooled eluate was dialyzed overnight at 4°C in buffer C (25 mM Tris pH 8.0, 150 mM NaCl) to remove imidazole. In the second step of the purification, the same HisTrap column was used to remove the cleaved His₆-SUMO fusion partner and the His₆-tagged SUMO protease. The resulting protein was >95% pure as judged by SDS-PAGE. A 10 ml HiTrap Desalting column (GE Life Sciences) was used to exchange the buffer to 25 mM bis-tris pH 7.1 (titrated with saturated iminodiacetic acid). The concentration of Swit_4259 was estimated from the absorbance at 280 nm (ϵ_{280} = 46 660 M⁻¹ cm⁻¹). The final yield of purified Swit_4259 was approximately 15 mg per litre of culture. Macromolecule-production information is summarized in Table 1.

2.2. Crystallization

Preliminary crystallization conditions were identified by sparse-matrix screening using the Index HT, PEG/Ion HT and PEGRx HT screens (Hampton Research, USA) at 21°C. Crystals appeared after 1–2 d in a variety of conditions. The

Table 2

Crystallization.

Method	Hanging-drop vapor diffusion
Plate type	24-well plate
Temperature (K)	295.0
Protein concentration (mg ml ⁻¹)	37
Buffer composition of protein solution	25 mM bis-tris pH 7.1 titrated with saturated iminodiacetic acid
Composition of reservoir solution	12% PEG 10 000, 0.1 M Tris pH 8.0, 0.3 M ammonium acetate
Volume and ratio of drop	2 µl; 1:1
Volume of reservoir (µl)	500

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	APS beamline 21-ID-F
Wavelength (Å)	0.97872
Temperature (K)	100
Detector	MAR Mosaic 225 mm CCD
Crystal-to-detector distance (mm)	200, 230
Rotation range per image (°)	0.5
Total rotation range (°)	100
Exposure time per image (s)	1, 0.5
Space group	<i>P</i> 3 ₂
<i>a</i> , <i>b</i> , <i>c</i> (Å)	75.04, 75.04, 179.85
α , β , γ (°)	90, 90, 120
Mosaicity (°)	0.48
Resolution range (Å)	44.96–1.91 (1.95–1.91)
Total No. of reflections	478983 (12577)
No. of unique reflections	87644 (4465)
Completeness (%)	100 (99.8)
Multiplicity	5.5 (2.8)
$\langle I/\sigma(I) \rangle$	6.8 (1.4)†
Mn(<i>I</i>) half-set correlation CC _{1/2}	0.993 (0.800)
<i>R</i> _{crim} ‡	0.117 (0.514)
Overall <i>B</i> factor from Wilson plot (Å ²)	32.77

† The resolution where the mean $\langle I/\sigma(I) \rangle$ falls below 2.0 is 2.00 Å. The CC_{1/2} and completeness indicate that a resolution of 1.91 Å is suitable for the data. ‡ Estimated $R_{\text{crim}} = R_{\text{merge}}[N/(N-1)]^{1/2}$, where *N* is the data multiplicity.

initial crystals ranged in shape from individual rods to what appeared to be bundles of parallel needles. The single, rod-shaped crystals diffracted very poorly ($d_{\text{min}} \simeq 5$ Å), but the ‘bundles of rods’ diffracted to a maximum resolution of ~2.0 Å. Diffraction-quality crystals were grown by the hanging-drop vapor-diffusion method, forming long, hexagonal rods of the order of 0.5 × 0.05 × 0.05 mm. Drops were formed by mixing 1–2 µl Swit_4259 (37 mg ml⁻¹) in 20 mM phosphate pH 7.0 with 1 µl well solution consisting of 12% PEG 10 000, 0.1 M Tris pH 8.0, 0.3 M ammonium acetate. Crystals were soaked in 10 mM 2-oxoadipate with the mother liquor and flash-cooled by plunging them into liquid nitrogen without a cryoprotectant. Crystallization information is summarized in Table 2.

2.3. Data collection and processing

Diffraction data were collected from an Swit_4259 crystal on the Life Sciences Collaborative Access Team beamline 21-ID-F at the Advanced Photon Source (APS) equipped with a MAR 225 CCD detector using a 50 × 50 µm beam at a wavelength of 0.97872 Å. A total of 400 frames were collected from $\omega = 10$ to 110° with an oscillation range of 0.5° (200

frames at 200 nm and 200 frames at 230 nm). Diffraction data were integrated and scaled using *MOSFLM* (Battye *et al.*, 2011) and *AIMLESS* (Evans & Murshudov, 2013) as implemented in the *CCP4* suite v.7.0 (Winn *et al.*, 2011). Data-collection and processing statistics are summarized in Table 3.

2.4. Structure solution and refinement

Swit_4259 crystallized with four molecules in the asymmetric unit. Initial phases were obtained by molecular replacement in *Phaser* (McCoy *et al.*, 2007) using the structure of the related ADCSF enzyme MppR (PDB entry 4jm3; 24.9% sequence identity; Burroughs *et al.*, 2013) as the starting model. Iterative rounds of model building and refinement

were performed using *PHENIX* (Adams *et al.*, 2010; Afonine, Mustyakimov *et al.*, 2010) and *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010). The structures were validated using the tools available in the *PHENIX* package (Afonine, Grosse-Kunstleve *et al.*, 2010; Chen *et al.*, 2010). The model-refinement statistics are summarized in Table 4.

3. Results and discussion

3.1. Overall structure of Swit_4259

Swit_4259 is a 28.8 kDa protein that, according to size-exclusion chromatography (Supporting Information §S1) and *PDBE*PISA assembly (Krissinel & Henrick, 2007), exists as a

Figure 2

C α -trace representations (a) of the *Chromobacterium violaceum* ADC (CvADC; PDB entry 3bgt; gray) structure superimposed on the Swit_4259 structure (blue) highlight the structural similarity between these two enzymes. The superimposition was performed using *SSM* (Krissinel & Henrick, 2004) in *Coot* (Emsley & Cowtan, 2004) and resulted in a root-mean-square deviation of 1.8 Å for 212 aligned *C* α atoms (of a total of 244 *C* α atoms in CvADC). The ribbon diagram (b) shows a single protomer and is in the same orientation as the orange protomer in (c). The main barrel of the β -cone fold is colored pale blue, with the 'side barrel' colored green. The dimer-interface loops are colored yellow, and the loops forming the interface between the dimers are colored orange. The section of the polypeptide that is colored red is the 'active-site loop' that may control access to the active site. The catalytic Lys122 is shown as a ball-and-stick representation with purple C atoms. The quaternary structure of Swit_4259 is shown as a solvent-accessible surface (c) and highlights the extensive tetramer interface formed by the long, interdigitating loops (top). This interface is repeated on the back side with the blue and pale blue protomers. The pair of dimers is held together (bottom) by short loops that appear to 'clip' the two dimers together to form the tetramer. This figure and subsequent figures showing crystallographic structures were rendered with the *POVScript+* (Fenn *et al.*, 2003) modification of *MolScript* (Kraulis, 1991) and *POV-Ray*.

Table 4

Structure solution and refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	44.06–1.91 (1.98–1.91)
Completeness (%)	99.1
σ Cutoff	2.0
No. of reflections, working set	167702 (16060)
No. of reflections, test set	5983 (561)
Final R_{cryst}	0.172 (0.2737)
Final R_{free}	0.201 (0.3071)
No. of non-H atoms	
Protein	7752
Ligand	0
Water	428
Total	8180
R.m.s. deviations	
Bonds (Å)	0.011
Angles (°)	1.176
Average B factors (Å ²)	
Overall	47.95
Protein	48.10
Water	45.25
Ramachandran plot	
Most favored (%)	99.11
Allowed (%)	0.89
No. of TLS groups	38

homotetramer in solution. Coordinates and structure factors have been deposited in the Protein Data Bank with accession code 5upb. As shown in Fig. 2(a), the overall fold of the Swit_4259 protomer is nearly identical to that of the proto-

typical family I ADCs, as well as the recently described family V enzymes such as SbAD from *Streptomyces bingchengensis* and MppR from *S. hygroscopicus* (not shown), which are not decarboxylases (Burroughs *et al.*, 2013; Mueller *et al.*, 2015). This fold, dubbed the β -cone (Ho *et al.*, 2009) or double-barrel fold (Tagaki & Westheimer, 1968), consists of a large, anti-parallel, 13-stranded β -sheet that forms the main barrel and folds back on itself to form a second, orthogonal barrel (Fig. 2b). Unlike prototypical ADCs that are active as homododecamers, Swit_4259 has a tetrameric quaternary structure similar to SbAD and MppR (Fig. 2c). The tetrameric architecture of Swit_4259 can be described as a dimer of dimers. There is a tight interface between chains *A* and *C* (pale yellow and orange protomers, respectively), and chains *B* and *D* (pale and dark blue protomers). These dimer units come together with a smaller interface where chain *A* contacts chain *B*, and chain *C* contacts chain *D*. The tetramer in Fig. 2(c) does not contain ligands, although the crystal was soaked in 10 mM 2-oxoadipate. The active sites from the four chains were similar, with two exceptions. One, chain *B*, contains electron density that is likely to correspond to polyethylene glycol from the crystallization cocktail in the active site; however, it is too indistinct to model with confidence. Secondly, chain *D* did not contain a second, stable water molecule (Wat2 in Fig. 3).

One important observation from all four active sites is that in the carboxylate-binding site one of the two key residues for

Figure 3

The residues constituting the carboxylate-binding site (Arg114 and Asn118) are different in Swit_4259 compared with the other family V ADCSF enzymes of known structure, SbAD and MppR, which both have an Arg–Gln pair. The electron density (a) shows that in contrast to SbAD and MppR, where Gln118 orients Arg114 to interact with the carboxyl group of pyruvate (and other α -keto acids) and holds it in position, the Asn118 residue in Swit_4259 does not interact with Arg114, which leads to a high degree of disorder in this region of the active site. Thus, it seems that while SbAD and MppR have active sites that are pre-organized to bind α -keto acids, Swit_4259 does not. This difference may reflect the proposed catabolic function of Swit_4259, given that MppR, at least, is known to be a biosynthetic enzyme (the biological role of SbAD is unknown). Magenta mesh, $2|F_o| - |F_c|$ map contoured at 2.0σ ; orange mesh, composite OMIT $2|F_o| - |F_c|$ map contoured at 1.0σ ; green and red mesh, $|F_o| - |F_c|$ map contoured at $+3.0$ and -3.0σ , respectively. The hydrogen-bonding networks and relevant interatomic distances (b) are shown for Swit_4259. Potential hydrogen-bonding interactions and their average distances in Å (from each of the four chains in the structure) are shown in green. All averages had standard deviations of 0.2 Å or less.

binding the carboxylate groups of α -keto acids in SbAD and MppR, Gln118, is replaced with Asn in Swit_4259. In SbAD and MppR, Gln118 orients Arg114 and holds it in position. In contrast, Asn118 of Swit_4259 points away from Arg114. Consequently, Arg114 is highly mobile as shown by its fragmented electron density (Fig. 3). The relatively malleable active site of Swit_4259 contrasts sharply with the active sites of SbAD and MppR (Figs. 4*a* and 4*b*, Supplementary Fig. S2 and Table S1), both of which possess carboxylate-binding sites within their active sites. MppR and SbAD can both be trapped crystallographically with a wide array of α -keto acids bound to their catalytic lysine residues. Numerous attempts to trap such complexes with Swit_4259 have been unsuccessful. It seems that whereas the active sites of MppR and SbAD, which are known or presumed to be biosynthetic enzymes, are pre-organized to bind pyruvate, the more disordered active site of Swit_4259 may reflect its putative catabolic function. Since its primary reaction is not likely to be carbon–carbon bond formation, there should be no requirement for a pre-ordered carboxylate-binding site to stabilize the enzyme–pyruvate complex until the second (aldehyde) substrate can bind and react.

3.2. Functional characterization

Swit_4259 did not demonstrate decarboxylase activity with acetoacetate (Supporting Information §S1), as no change at

270 nm was observed (Highbarger *et al.*, 1996). The high degree of structural similarity between the MppR, SbAD and Swit_4259 active sites suggested that the latter enzyme would also accept an α -keto acid as a substrate. Our previous work on SbAD found that Tyr252 plays a key role in the aldolase-dehydratase activity of this enzyme (Mueller *et al.*, 2015). Likewise, the equivalent residue in MppR, Glu283, appears to be important for its cyclization reaction (Burroughs *et al.*, 2013). The amino acid at this position appears to be a key determinant of the reaction and/or substrate specificity of ADCSF enzymes. The residue occupying this position in Swit_4259 is Arg205. The occurrence of arginine at this position suggested that Swit_4259 might bind an α -keto dicarboxylic acid. For this reason, 2-oxo-hept-3-enedioate (**1**; Fig. 5*a*) was qualitatively modelled into the Swit_4259 active site (Fig. 6*a*) based on an overlay with the pyruvate-bound structure of SbAD (PDB entry 4zbt; Mueller *et al.*, 2015). While not based on experimental data, this model did suggest that α -keto dicarboxylic acids of sufficient length are likely to bind to the enzyme.

The binding constants of pyruvate, 2-oxoglutarate and 2-oxoadipate were measured by fluorescence titration (Fig. 6*b*, Table 5 and Supporting Information §S1). The K_d of pyruvate was comparable to that measured for SbAD (Mueller *et al.*, 2015), and in the realm of estimates of the intracellular pyruvate concentration in bacteria (Roy & Packard, 1998). Oxoglutarate had the same K_d as pyruvate, suggesting that its

Figure 4

Orthogonal views of overlays of the wild-type Swit_4259 (gold; PDB entry 5upb), SbAD (light blue; PDB entry 4zbo) and MppR (dark blue; PDB entry 4jm3) structures are shown in (a) and (b); Glu84, Val106 and Ser111 have been omitted for clarity in (b). Residue labels refer to Swit_4259; the equivalent residues in SbAD and MppR have been omitted for clarity. The carboxylate-binding site residues of SbAD and MppR (primarily Arg114 and Gln118) form a well ordered binding site for pyruvate that does not change significantly upon pyruvate binding. The 'carboxylate-binding site' of Swit_4259 has Asn at position 118 rather than Gln, and this change appears to disrupt the carboxylate-binding site. While Arg114 is in approximately the same position in Swit_4259 as observed in the other two proteins, the *B* factors are much higher and there is evidence of alternative conformations for the Arg side chain. Both of these observations indicate increased mobility of this region of the Swit_4259 active site relative to the other family V enzymes.

side chain is too short to take advantage of the carboxylate–Arg205 interaction. 2-Oxo-adipate had a somewhat lower K_d ,

suggesting that the longer, aliphatic side chain positions the carboxylate group to interact with Arg205.

Figure 5

(a) 2-Oxo-hept-3-enedioate (**1**) is postulated to be a substrate of Swit_4259. (b) 2-Oxo-hex-3-enedioate (**4**) was produced by treatment of 2-hydroxymuconic acid (**3**) with 4-oxalocrotonate tautomerase (4OT). The aldol product **5** was observed after treatment with Swit_4259; the retro-aldol cleavage products malonic semialdehyde (**6**) and pyruvate (**7**) were not. (c) Vinyl pyruvate (**8**) incubated with Swit_4259 produced the aldol product (**9**). (d) Swit_4259-catalyzed aldol condensation of 3-(2-furyl)acrolein **10** and pyruvate (**7**) to give 3-(2-furyl)acroleinpyruvate (**11**).

Figure 6

(a) The volume of the active-site cavity in Swit_4259 calculated by VOIDOO (Kleywegt & Jones, 1994; Kleywegt *et al.*, 2001) is shown as a green mesh. The hypothetical ligand 2-oxo-hept-3-enedioate (transparent spheres) was qualitatively modeled in the active site based on the pyruvate-bound structure of the related enzyme SbAD. Notice that the length of this ligand permits the carboxylate group of the side chain to interact with Arg205. (b) The binding constants for pyruvate (squares), α -ketoglutarate (triangles) and 2-oxoadipate (circles) were determined by fluorescence titration of Swit_4259 (see Table 5 for the derived dissociation constants). Traces have been offset by 0.25 units along the ordinate axis for clarity.

Table 5

Binding constants for the interaction of Swit_4259 with three α -keto acids.

Substrate	K_d (mM)
Pyruvate	10.1 ± 0.9
2-Oxoglutarate	11.0 ± 1.4
2-Oxoadipate	5.9 ± 0.7

Given the binding constant of 2-oxoadipate, we postulated that the seven-carbon **1** (one carbon unit longer than 2-oxoadipate) would be a viable substrate for Swit_4259. Since **1** was not available for kinetic tests, we generated a similar compound, 2-oxo-hex-3-enedioate (**4**; Fig. 5*b*), from 2-hydroxy-muconic acid (**3**) using the enzyme 4-oxalocrotonate tautomerase (4OT), a kind gift from the Whitman group at the University of Texas at Austin. When $100 \mu\text{M}$ **4** was incubated with $10 \mu\text{M}$ Swit_4259, there was a steady decrease in the absorbance at 236 nm owing to the loss of conjugation of **4** (Fig. 7*a*). The rate of decrease was dependent on both the concentration of **4** and the concentration of the enzyme (data not shown).

In order to identify the product of the reaction of Swit_4259 with **4**, reaction mixtures were analysed by ESI-MS and ^1H NMR. The ESI-MS data did not show a signal at $m/z = 87$ that would indicate the presence of the retro-aldol cleavage products malonic semialdehyde (**6**) and pyruvate (**7**) in the reaction. The only relevant signal that was observed, at $m/z = 174$, is consistent with the aldol product **5** (Fig. 5*b*). Vinyl pyruvate (**8**; Fig. 5*c*) was also treated with Swit_4259 and the reaction mixture was analyzed by ESI-MS and ^1H NMR.

As with **4**, only the aldol product could be detected; there was no evidence of pyruvate in the reaction mixture. There was no evidence of the aldol forms **5** (Fig. 5*b*) or **9** (Fig. 5*c*) in standards that had not been incubated with Swit_4259. The NMR spectra for reactions with both substrates, **4** (Supplementary Figs. S5 and S6) and **8**, showed that a mixture of compounds were present, but were too complicated to interpret unambiguously. However, it was clear from these spectra that there was not a detectable amount of pyruvate present (Supplementary Fig. S5), as there was not a significant peak at the resonance of the methyl H atoms in the pyruvate standard at 2.1 p.p.m.

The steady-state kinetics of the reaction of Swit_4259 with **4** were analysed for both the wild-type enzyme and the Asn118Glu variant (Table 6). Asn118 was mutated to Glu in an effort to recapitulate the more ordered carboxylate-binding sites of SbAD and MppR (Figs. 4*a* and 4*b*, Supplementary Fig. S2 and Table S1). The more malleable active site of Swit_4259, which leads to subtle differences in the active sites of pairs of protomers within the tetramer, suggests that cooperativity could be possible, but our preliminary steady-state kinetics do not reflect any evidence of cooperativity (Table 6). Based on the data in Table 3, it is clear that **4** is not the physiological substrate of Swit_4259. The K_m value suggests that **4** binds reasonably well to the enzyme, but the turnover number is quite low. It is likely that **4** does not bind in an orientation suitable to promote high catalytic efficiency. A seven-carbon α -keto acid such as 2-oxo-hept-3-ene-dioate (**1**), rather than the six-carbon **4**, could conceivably fill the active site more completely and be a more efficient substrate. This is supported

Figure 7

UV-visible spectra showing the Swit-4259-catalyzed breakdown of **4** (*a*). A solution of $100 \mu\text{M}$ **3** was treated with $1 \mu\text{M}$ 4-oxalocrotonate tautomerase for 20 min before Swit_4259 was added to a final concentration of $10 \mu\text{M}$. Scans were recorded every 2 min at 25°C and are shown colored from black ($t = 0$) to magenta ($t = 44$); every other scan has been omitted for clarity. The absorbance at 300 nm is owing to residual **3**, while the peak at 236 nm is owing to **4**. Reaction mixtures of Swit_4259 with pyruvate and **10** were analyzed by reverse-phase HPLC (*b*) and show that Swit_4259 is able to condense pyruvate and **10**, but appears to be incompetent to perform the reverse reaction. The mixture of $500 \mu\text{M}$ **10** with 50 mM pyruvate (red trace) shows that under the conditions of the separation **10** has a retention time of 4.0 min. A standard of synthetically prepared **11** shows that this compound has a retention time of 5.6 min (not shown). When Swit_4259 was added to the mixture of **10** and pyruvate and allowed to stand for 2 h at RT (blue trace), all of the **10** was converted to **11**. Conversely, when Swit_4259 was added to a sample of pure **11**, no **10** was produced, even after 16 h at RT.

Table 6

Steady-state kinetic parameters for the reaction of Swit_4259 with **4** or with 50 mM pyruvate and **10**.

Substrate	Swit_4259 WT	Swit_4259 Asn118Glu
4 (Fig. 5 <i>b</i>)		
$k_{\text{cat},4}$ (s^{-1})	0.010 ± 0.001	No activity
$K_{\text{m},4}$ (μM)	66.4 ± 16	
$k_{\text{cat},4}/K_{\text{m},4}$ ($\text{M}^{-1} \text{s}^{-1}$)	$(1.5 \pm 0.4) \times 10^2$	
10 (Fig. 5 <i>d</i>)		
$k_{\text{cat},10}$ (s^{-1})	0.08 ± 0.01	0.07 ± 0.01
$K_{\text{m},10}$ (mM)	3.0 ± 0.7	3.6 ± 0.8
$k_{\text{cat},10}/K_{\text{m},10}$ ($\text{M}^{-1} \text{s}^{-1}$)	26.7 ± 7.1	19.4 ± 5.1

by steady-state kinetic analysis of the Swit_4259-catalyzed aldol condensation of 3-(2-furyl)acrolein (**10**; Fig. 5*d*) and pyruvate (Table 6), which approximates a seven-carbon long substrate.

We chose to look at the reaction of Swit_4259 with **10** based on the close relationship between Swit_4259 and SbAD, which was observed to catalyze the reaction shown in Fig. 5(*d*) (Mueller *et al.*, 2015). Wild-type (WT) Swit_4259 and the Asn118Glu variant were both able to catalyze the aldol condensation with turnover numbers that were sevenfold to eightfold faster than the reaction with **4**. The activity of the Asn118Glu variant was nearly identical to that of the WT. The product of the Swit_4259-catalyzed aldol condensation was confirmed by reverse-phase HPLC (Fig. 7*b*), which showed that the product of the reaction matched the retention time of a standard sample of **11**. We were surprised to find that neither WT Swit_4259 nor the Asn118Glu variant catalyzed the retro-aldol cleavage of **11**. Neither the aldol nor pyruvate products were detectable by HPLC (Fig. 7*b*), ESI-MS or ^1H NMR. It is possible that the high K_{m} value for **10** reflects poor affinity for this substrate, and the affinity of the enzyme for the product, **11**, may be yet lower. The activity of Swit_4259 is certainly distinct from those of the classical ADCs, as well as the family V enzymes SbAD, which is an aldolase dehydratase that can catalyze both the condensation and cleavage reactions (Mueller *et al.*, 2015), and MppR, which catalyzes the cyclization of an oxidized arginine derivative (Burroughs *et al.*, 2013; Supplementary Fig. S1).

4. Conclusion

The clear differences in the dynamics of the active sites comprising the two dimers of the tetramer imply that Swit_4259 could be a cooperative enzyme. However, thus far our preliminary kinetics studies show no evidence of cooperativity. While Swit_4259 exhibits very weak activity with the substrates tested, it can catalyze the aldol condensation and subsequent dehydration of pyruvate and 3-(2-furyl)acrolein to yield the corresponding enone; the enzyme was not able to catalyse the reverse reaction. ESI-MS analysis shows that both **4**, which is a better substrate in terms of its ability to bind to the enzyme, and vinyl pyruvate (**8**) are only hydrated by Swit_4259. Future work will focus on characterizing the other members of the proposed catabolic gene cluster, with

the goal of identifying the physiological substrate for this putative catabolic pathway.

5. Related literature

The following references are cited in the Supporting Information for this article: Metanis *et al.* (2005), Robert & Gouet (2014) and Sievers *et al.* (2011).

Acknowledgements

We thank Dr Christian P. Whitman for the gifts of 2-hydroxymuconic acid and the expression plasmid for 4-oxalocrotonate tautomerase. The authors declare no competing financial interests. Author contributions are as follows. LSM expressed, purified and crystallized Swit_4259 and conducted steady-state kinetics experiments. ZM solved the initial Swit_4259 crystal structure and performed end-point assays. TWK expressed and purified Swit_4259 and crystallized the enzyme. TF synthesized 2-hydroxymuconic acid to support the steady-state kinetics experiments. TH and RWH produced 3-(2-furyl)acroleinpyruvate. AWS conceived of and supervised the organic synthesis work. TRM was vital to discussions of the potential function of Swit_4259. NRS conceived the idea for the project and prepared the manuscript with LSM. We thank Dr Graham R. Moran for looking over the manuscript.

Funding information

The following funding is acknowledged: National Science Foundation, Directorate for Biological Sciences (grant No. MCB-1157392 to Nicholas R. Silvaggi). This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817).

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Afonine, P. V., Grosse-Kunstleve, R. W., Chen, V. B., Headd, J. J., Moriarty, N. W., Richardson, J. S., Richardson, D. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2010). *J. Appl. Cryst.* **43**, 669–676.
- Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R. W., Moriarty, N. W., Langan, P. & Adams, P. D. (2010). *Acta Cryst.* **D66**, 1153–1163.
- Aly, H. A. H., Huu, N. B., Wray, V., Junca, H. & Pieper, D. H. (2008). *Appl. Environ. Microbiol.* **74**, 3812–3822.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denancé, N., Vasse, J., Lauber, E. & Arlat, M. (2007). *PLoS One*, **2**, e224.
- Burroughs, A. M., Hoppe, R. W., Goebel, N. C., Sayyed, B. H., Voegtline, T. J., Schwabacher, A. W., Zabriskie, T. M. & Silvaggi, N. R. (2013). *Biochemistry*, **52**, 4492–4506.
- Chai, B., Tsoi, T. V., Iwai, S., Liu, C., Fish, J. A., Gu, C., Johnson, T. A., Zylstra, G., Teppen, B. J., Li, H., Hashsham, S. A., Boyd, S. A., Cole, J. R. & Tiedje, J. M. (2016). *PLoS One*, **11**, e0157008.

- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Colquhoun, D. R., Hartmann, E. M. & Halden, R. U. (2012). *J. Biomed. Biotechnol.* **2012**, 408690.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204–1214.
- Fenn, T. D., Ringe, D. & Petsko, G. A. (2003). *J. Appl. Cryst.* **36**, 944–947.
- Halden, R. U. & Dwyer, D. F. (1997). *Biorem. J.* **1**, 11–25.
- Highbarger, L. A., Gerlt, J. A. & Kenyon, G. L. (1996). *Biochemistry*, **35**, 41–46.
- Ho, M.-C., Ménétret, J.-F., Tsuruta, H. & Allen, K. N. (2009). *Nature (London)*, **459**, 393–397.
- Kleywegt, G. J. & Jones, T. A. (1994). *Acta Cryst.* **D50**, 178–185.
- Kleywegt, G. J., Zou, J.-Y., Kjeldgaard, M. & Jones, T. A. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 353–356. Dordrecht: Kluwer Academic Publishers.
- Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
- Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* **D60**, 2256–2268.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Li, Q., Wang, X., Yin, G., Gai, Z., Tang, H., Ma, C., Deng, Z. & Xu, P. (2009). *Environ. Sci. Technol.* **43**, 8635–8642.
- Magarvey, N. A., Haltli, B., He, M., Greenstein, M. & Hucul, J. A. (2006). *Antimicrob. Agents Chemother.* **50**, 2167–2177.
- Malakhov, M. P., Mattern, M. R., Malakhova, O. A., Drinker, M., Weeks, S. D. & Butt, T. R. (2004). *J. Struct. Funct. Genomics*, **5**, 75–86.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Metanis, N., Keinan, E. & Dawson, P. E. (2005). *J. Am. Chem. Soc.* **127**, 5862–5868.
- Miller, T. R., Delcher, A. L., Salzberg, S. L., Saunders, E., Detter, J. C. & Halden, R. U. (2010). *J. Bacteriol.* **192**, 6101–6102.
- Monna, L., Omori, T. & Kodama, T. (1993). *Appl. Environ. Microbiol.* **59**, 285–289.
- Mueller, L. S., Hoppe, R. W., Ochsenwald, J. M., Berndt, R. T., Severin, G. B., Schwabacher, A. W. & Silvaggi, N. R. (2015). *Biochemistry*, **54**, 3978–3988.
- Nam, I.-H., Hong, H.-B. & Schmidt, S. (2014). *J. Microbiol.* **52**, 801–804.
- Nguyen, T. P. O., Helbling, D. E., Bers, K., Fida, T. T., Wattiez, R., Kohler, H.-P. E., Springael, D. & De Mot, R. (2014). *Appl. Microbiol. Biotechnol.* **98**, 8235–8252.
- Postle, K. & Kadner, R. J. (2003). *Mol. Microbiol.* **49**, 869–882.
- Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
- Roy, S. O. & Packard, T. T. (1998). *Appl. Environ. Microbiol.* **64**, 4958–4964.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol. Syst. Biol.* **7**, 539.
- Tagaki, W. & Westheimer, F. H. (1968). *Biochemistry*, **7**, 895–900.
- Tang, K., Jiao, N., Liu, K., Zhang, Y. & Li, S. (2012). *PLoS One*, **7**, e41204.
- Wang, H., Xiao, Y., Fu, L., Zhao, H., Zhang, Y., Wan, X., Qin, Y., Huang, Y., Gao, H. & Li, X. (2010). *BMC Biotechnol.* **10**, 14.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Wittich, R. M., Wilkes, H., Sinnwell, V., Francke, W. & Fortnagel, P. (1992). *Appl. Environ. Microbiol.* **58**, 1005–1010.