

Engineering a more specific *E. coli* glyoxylate/hydroxypyruvate reductase for coupled steady state kinetics assays

Nemanja Vuksanovic^a, Dante A. Serrano^a, Brandon M. Patterson^a, Nicholas R. Silvaggi^{a}*

^aDepartment of Chemistry and Biochemistry, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA

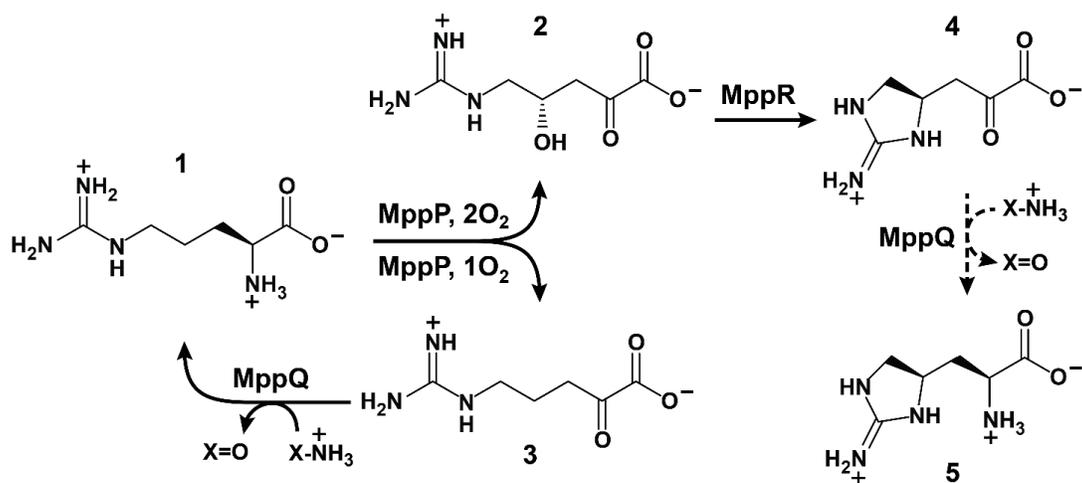
FUNDING STATEMENT: This work was supported by grant CHE-1903899 from the National Science Foundation, Division of Chemistry.

KEYWORDS: Glyoxylate reductase, 2-ketoarginine, protein engineering, selectivity, coupling enzyme

ABBREVIATIONS: GLX, glyoxylate; 2KA, 2-ketoarginine; EcGhrA, *E. coli* glyoxylate reductase; L-End, L-enduracididine, OA, oxaloacetate, α KG, α -ketoglutarate

ABSTRACT: The *E. coli* glyoxylate reductase/hydroxypyruvate reductase A (EcGhrA) was investigated as a coupling enzyme to monitor the transamination of 2-ketoarginine and glycine by the L-enduracididine biosynthetic enzyme MppQ. Surprisingly, 2-ketoarginine proved to be an efficient substrate for EcGhrA. Since the promiscuity of EcGhrA prevented its use as a coupling enzyme to monitor the aminotransferase activity of MppQ, we set about engineering a more specific variant. X-ray crystal structures of EcGhrA were determined in the unliganded state, as well as with glyoxylate and 2-ketoarginine bound. The electron density maps of EcGhrA with 2-ketoarginine bound showed weak electron density for the side chain of this substrate, complicating the choice of active site residues to target for site-directed mutagenesis. The structure of the complex did, however, suggest that the side chain of W45 could interact with the guanidinium group of 2-ketoarginine. We therefore generated the EcGhrA^{W45F} variant and tested it for activity with 2-ketoarginine, glyoxylate, oxaloacetate, α -ketoglutarate, α -oxofuranacetic acid, phenyl pyruvate, 3-mercaptopyruvate and 2-ketobutyric acid. The W45F variant exhibited a ~10-fold decrease in the specificity constant (k_{cat}/K_M) for 2-ketoarginine, while the reaction with glyoxylate was not significantly impaired. The reactions of the W45F variant with the alternative substrates oxaloacetate and α -ketoglutarate were also impaired. Thus, the W45F variant is a less promiscuous enzyme than the wild-type. This engineered EcGhrA^{W45F} variant could be generally useful as a coupling system for enzymes that produce glyoxylate, such as 4-hydroxy-2-oxoglutarate aldolase or isocitrate lyase.

The non-proteinogenic amino acid L-enduracididine (Scheme 1, **5**) is a component of several non-ribosomally produced peptide natural products, including the antibiotic mannopeptimycin^{1, 2}. As a part of our work on L-enduracididine biosynthesis,³⁻⁵ it became necessary to develop a coupled assay to measure the kinetics of the putative 2-ketoenduracididine aminotransferase MppQ. In addition to its proposed role in the last step of L-enduracididine biosynthesis, MppQ is also capable of returning the dead-end product 2-ketoarginine (Scheme 1, **3**) back to the starting material, L-Arg. Preliminary experiments showed that, in reactions with L-Arg and a selection of amino acceptor substrates (*e.g.* glyoxylate, pyruvate, oxaloacetate and α -ketoglutarate, Scheme 2), MppQ was most efficient using glyoxylate as the acceptor substrate (unpublished data). This led to the idea that, in the more biosynthetically relevant transamination of 2-ketoenduracididine or 2-ketoarginine to the corresponding amino acids, glycine should be an efficient amino donor substrate. Since these reactions would result in the production of glyoxylate, we selected GhrA from *E. coli* (EcGhrA) as a potential coupling enzyme in a coupled spectrophotometric assay for

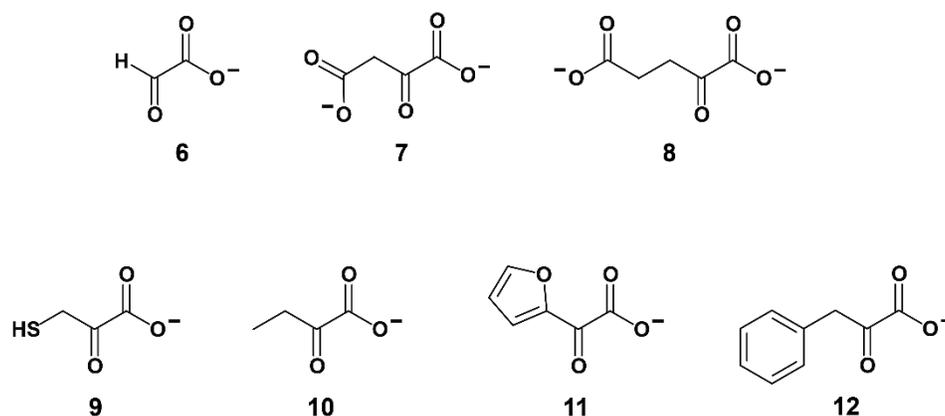


Scheme 1. The L-enduracididine biosynthetic pathway begins with L-arginine (**1**), which is oxidized by the unusual PLP-dependent L-Arg oxidase, MppP, to 4(S)-hydroxy-2-ketoarginine (**2**). MppP, at least *in vitro*, releases the partially oxidized dead-end product 2-ketoarginine (**3**) in about 40% of catalytic cycles. The oxidized arginine derivative **2** is cyclized by MppR to create the iminoimidazolidine ring of the intermediate **4** (*i.e.* 2-ketoenduracididine). MppQ is thought to catalyze the transamination of **4** with an as-yet-unidentified amino donor substrate to yield the finished amino acid L-enduracididine (**5**). MppQ also catalyzes the transamination of the dead-end product **3** to regenerate L-Arg, though it has not been ascertained if this is a physiologically relevant catalytic activity or an adventitious *in vitro* phenomenon.

the transamination of 2-ketoarginine or 2-ketoenduracididine and glycine. This particular enzyme was chosen over other enzymes previously used in assays of glyoxylate-producing reactions, such as porcine L-lactate dehydrogenase⁶ or the glyoxylate reductase from *Spinacea oleracea*,⁷ because it could be produced very efficiently and inexpensively in-house.

The D-2-hydroxyacid dehydrogenases (2HAD) family contains enzymes with diverse substrate specificities that show significant potential for use in the industrial synthesis of chiral D-2-hydroxy acids.⁸ Members of the 2HAD family are known to be promiscuous, and EcGhrA also catalyzes the reduction of 2-hydroxypyruvate.⁹ Surprisingly (at least to us), the promiscuity of EcGhrA extended to the much larger and positively charged 2-ketoarginine, which would, of course, preclude its use in monitoring the reaction we wished to study.

This observation led us to wonder what other α -keto acids are substrates for EcGhrA. The substrate specificities of other GhrA homologs have been examined and it was found that oxaloacetate is a substrate for the *Euglena gracilis*¹⁰ and *Saccharomyces cerevisiae*¹¹ enzymes, succinic semialdehyde is a substrate for the *Arabidopsis thaliana* homolog¹², *Neurospora crassa* GhrA reacts with phenylpyruvate,¹³ and α -ketoglutarate is a slow substrate for the *E. gracilis* GhrA¹⁰. To extend the knowledge of EcGhrA promiscuity, we investigated additional α -keto acids (Scheme 2, 7-12), some of which have not been tested previously. The results of these studies,



Scheme 2. EcGhrA substrates investigated in this work.

together with the structural work described herein, suggested that it might be possible engineer an EcGhrA variant with increased specificity for glyoxylate in order to develop a suitable coupling enzyme for use in future steady state kinetics work.

Rational engineering of 2HAD enzymes to alter substrate specificity has been described in the literature. In 2007, Shinoda et al. transformed formate dehydrogenase into a glyoxylate reductase by mutating two catalytic residues¹⁴. In 2013, Zhen et al. rationally engineered D-lactate dehydrogenase to reduce phenylpyruvic acid, α -ketobutyric acid, α -ketovaleric acid, and α -hydroxypyruvic acid via a single Y52L mutation.¹⁵ At the time of this writing, rational engineering of a GhrA homolog to make it more specific for glyoxylate has not been attempted.

Herein we describe the crystal structures of wild-type EcGhrA in ternary complexes with NADP⁺ and either glyoxylate or **3** that were the basis of our (ir)rational design of a more specific glyoxylate reductase. These structures suggested that making a limited number of mutations near the outer edge of the active site could reduce the affinity of EcGhrA for **3**. Steady state kinetics of the wild-type and variant EcGhrA enzymes shows that a single mutation, W45F, leads to a 10-fold decrease in the specificity constant (k_{cat}/K_M) for **3** without significant impact on the reaction with glyoxylate. This new variant of EcGhrA may be useful for studying the kinetics of glyoxylate-producing enzymes where the promiscuity of the wild-type EcGhrA is problematic.

Materials and Methods

Cloning, expression and purification of wild-type and mutant EcGhrA. The EcGhrA gene was amplified from *E. coli* genomic DNA using primers containing NdeI and BamHI restriction sites (underlined; forward, 5'-GGATTCCCATATGGAAAACCTGTATTTTCAGGGTATGGATATCATCTTTTATCACCCAA-3'; reverse, 5'-CGGGATCCTTAGTAGCCGCGTGCGC-3').

The forward primer also encoded a tobacco etch virus (TEV) protease cleavage site (*italic*, translates to MENLYFQG). The PCR reaction was carried out using Phusion High Fidelity PCR Master Mix (Thermo Scientific) and the genomic DNA was extracted from *E. coli* BL21 Star (DE3) (Invitrogen) using genomic the Blood and Cell Culture DNA mini kit (Qiagen). The amplified PCR product was digested in one step using NdeI and BamHI HF (New England Biolabs) and ligated into similarly treated pET-15b to create plasmid pET15b-EcGhrA. Site-directed mutagenesis to create the EcGhrA^{W45F} variant was performed using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) using the mutagenic primers 5'-TGCTTTAGTCTTTCATCCTCCTG-3' (forward) and 5'-TAATCAGCAGAGTCATTATC-3' (reverse).

The His₆-tagged wild-type EcGhrA was expressed from *E. coli* BL21 Star (DE3) cells carrying the pET15b-EcGhrA plasmid grown at 37 °C in Luria-Bertani broth with 100 µg/mL ampicillin. Protein expression was induced with 0.4 mM IPTG when the OD_{600nm} reached 0.6. The temperature was reduced to 25 °C and the cells were grown overnight with shaking at 250 rpm. Cells were harvested via centrifugation and resuspended in 5 mL of Buffer A [25 mM TRIS (pH 8.0), 300 mM NaCl, 10 mM imidazole] containing 0.1 mg/mL Dnase I (Worthington Biochemical Corp.) per gram of wet cell paste. Cells were lysed using a Branson Sonifier S-450 cell disruptor (Branson Ultrasonics Corp.) at 60% amplitude with 30 s pulses separated by 50 s rest periods for a total of 10 min of sonication. The temperature of the lysate was maintained below 4 °C by keeping the steel beaker suspended in an ice bath over a stir bar. The lysate was clarified by centrifugation at 39,000 x g for 45 min before loading it onto a 5 mL HisTrap column (GE Lifesciences) at a flow rate of 5 mL/min. The His₆-tagged protein was purified using a 4-step gradient of increasing concentrations of Buffer B [25 mM TRIS (pH 8.0), 300 mM NaCl, and 250 mM imidazole; 5 column volumes each at 5, 15, 50, and 100% B]. The elution was monitored spectroscopically at 280 nm. His₆-

EcGhrA eluted at 50% Buffer B, and these fractions were collected and analysed using Coomassie-stained SDS-PAGE. In order to remove the His₆ tag, the pooled fractions were dialyzed overnight against 10 μM TEVpM2 (an optimized TEV protease variant)¹⁶ in 50 mM TRIS (pH 8.0), 100 mM NaCl, 1 mM DTT, and 0.25 mM EDTA at 4 °C. The dialysate was passed through the same HisTrap column, but less than 10% of the protein had the His₆ tag removed. The uncleaved protein was desalted into 10 mM HEPES pH 7.5, 1 mM DTT, by dialyzing overnight at 4 °C before being aliquoted and stored at -80 °C. EcGhrA^{W45F} was expressed and purified in the same manner as the wild type EcGhrA.

Synthesis of 2-ketoarginine, 3: Compound **3** was synthesized by enzymatic oxidation of L-arginine as described by Meister¹⁷ and purified as follows. Residual L-arginine was removed by passing the mixture through Dowex 50 WX8 (Sigma Aldrich) that had been equilibrated with approximately 10 CV of 0.5 M HCl, followed by 10 CV of 0.5 M NaCl, and finally 10 CV of water¹⁸. The identity and purity of **3** were confirmed by NMR spectroscopy (see Supplementary Information, Figure S1).

Steady state enzyme kinetics: The initial velocities of the wild-type and mutant EcGhrA-catalyzed reductions of varied concentrations of glyoxylate, and concomitant oxidation of NADPH, under steady state conditions were monitored at 340 nm ($\epsilon_{340\text{nm}} = 6220 \text{ M}^{-1}\text{cm}^{-1}$) using a Shimadzu UV-2600 double beam UV-visible spectrophotometer. All the reactions were carried out in triplicate under identical conditions. The enzyme (0.5 μM) was incubated with NADPH (200 μM) for 30 seconds in 50 mM HEPES, pH 7.5 at 22 °C before initiating the reaction by the addition of substrate. The substrate concentrations ranged from 0.2 to 40 mM for **6** and **3**, and 0.6 to 40 mM for **7** and **8**. Compounds **9-12** were tested only at 40 mM. No reactions were observed with these substrates, so they were not tested at lower concentrations. The steady state kinetic constants for

the reactions with evidence of substrate inhibition (e.g. wild-type EcGhrA with **6**) were estimated by fitting the initial velocity data to Equation 1:

Equation 1
$$v_o = \frac{V_{\max}S}{K_M + S(1 + \frac{S}{K_i})}$$

where v_o is the initial velocity, V_{\max} is the maximum velocity, S is the substrate concentration, K_M is the Michaelis constant, and K_i is the inhibition constant for reversible substrate inhibition. Reactions showing no evidence of substrate inhibition were analyzed using the standard Michaelis-Menten equation (Equation 2):

Equation 2
$$v_o = \frac{V_{\max}S}{K_M + S}$$

Crystallization of the EcGhrA·NADP⁺ binary complex: Diffraction quality crystals were obtained after screening ~16 mg/mL EcGhrA against the Index HT screen (Hampton Research). The optimized crystallization buffer consisted of 1.26 M sodium phosphate monobasic monohydrate, 0.14 M potassium phosphate dibasic, pH 5.6, and 1.8 mM NADP⁺. Hanging drop vapor diffusion experiments were initiated by mixing 1 μ L of ~16 mg/mL wild-type EcGhrA in 10 mM HEPES pH 7.5, 1 mM DTT with 1 μ L of the crystallization solution on a siliconized glass cover slip and sealing this above a well containing 500 μ L of the crystallization solution in a 24-well VDX plate (Hampton Research). Hexagonal prisms of approximate dimensions 200 x 50 x 50 μ m formed after 1 day. For the binary complex, crystals were cryo-protected by a brief soak in the crystallization buffer supplemented with 20 % glycerol.

*Crystallization of the ternary complexes of EcGhrA with NADP⁺ and **3**, **6**, or **8**:* Since we observed that phosphate significantly inhibits EcGhrA activity at the concentrations used for crystallization (data not shown) and is therefore likely to compete with the substrate for binding at the active site, we identified an alternative, phosphate-free crystallization condition. After

optimization of the initial condition from the Index HT screen, diffraction-quality crystals were obtained from hanging drop vapor diffusion experiments where the well solution consisted of 17% PEG 3350, 0.2 lithium sulfate, 0.1 M BIS-TRIS, pH 5.5, and 1.8 mM NADP⁺. The crystals were soaked overnight in 10 μ L of a solution containing 50 mM substrate (**3**, **6**, or **8**), 5 mM NADP⁺, 30% PEG 3350, 0.2 M lithium sulfate, and 0.1 M BIS-TRIS, pH 5.5. Paratone N (Hampton Research) was used to cryo-protect the crystals before being flash-cooled by plunging in liquid nitrogen.

Crystallization of EcGhrA^{W45F} with NADP⁺: Diffraction-quality crystals were obtained from the LMB screen (Molecular Dimensions) at 295 K, in a sitting drop vapor diffusion experiment where the well solution consisted of 16% PEG 4000, 0.1 ammonium sulfate, 0.1 M sodium citrate, pH 5.8, 20% glycerol and 1 mM NADP⁺. The crystal used for data collection was taken directly out of the drop and flash-cooled by plunging in liquid nitrogen. No further cryo-protection was required owing to the 20% glycerol present in the crystallization mixture.

Data collection and structure solution: X-Ray diffraction data for the EcGhrA·NADP⁺ binary complex and the EcGhrA·NADP⁺·**3** ternary complex were collected using the Life Sciences Collaborative Access Team (LS-CAT) beamline 21-ID-F at the Advanced Photon Source (APS). Diffraction data for the EcGhrA·NADP⁺·**6** and EcGhrA·NADP⁺·**8** ternary complexes were collected using LS-CAT beamline 21-ID-G. Both beamlines were equipped with MAR 300 CCD detectors and 50 x 50 μ m beams at wavelengths of 0.97856 Å and 0.97872 Å, respectively. In all cases, data were collected for 180 ° sweeps of the crystals with an oscillation range of 0.5° for each image. The exposure time was 1.0 s per image (2.0 s per °). The crystal-to-detector distances for the EcGhrA·NADP⁺ binary complex and the three ternary complexes ranged from 250 to 300 mm. The data for the EcGhrA^{W45F}·NADP⁺ binary complex was collected on LS-CAT beamline 21-ID-

D. This beamline was equipped with a Dectris Eiger 9M detector and a 50 x 50 μm beam at a wavelength of 1.1272 \AA . A total of 900 frames were collected from $\varphi = 0$ to 180 $^\circ$ at a scan rate of 5 $^\circ/\text{sec}$, with 5 images collected per degree, at a detector distance of 200 mm. All X-ray diffraction data sets were indexed, integrated and scaled using HKL2000¹⁹.

Initial phase estimates for all five structures were determined by molecular replacement in PHASER²⁰ as implemented in the CCP4²¹ suite. The structure of the binary complex of EcGhrA with NADP⁺ was determined using the coordinates of chain A of GhrA from *Salmonella typhimurium* (PDB ID 3KBO, unpublished) as the search model, with all non-protein atoms removed and all isotropic B-factors set to 20 \AA^2 . After refinement of this model (see below), it was used as the search model for phasing the three ternary complex structures and the EcGhrA^{W45F} structure. The EcGhrA·NADP⁺ structure was prepared for molecular replacement by removing all hydrogen atoms, solvent molecules, and NADP⁺ and setting all isotropic B-factors to 20.0 \AA^2 . Model refinement was performed using phenix.refine from the PHENIX suite^{22, 23} and model building was done with COOT^{24, 25}. Iterative rounds of model building and refinement were done until the R-factors converged, at which point the substrates were modeled into the ternary complex structures. The optimal number of translation-libration-screw (TLS) groups was determined using phenix.find_tls_groups (P. V. Afonine, unpublished). The final models were validated using MolProbity²⁶, as implemented in the PHENIX suite. Residues lacking defined electron density for the sidechain were modeled without the sidechains in order to avoid interpretational bias. The geometric restraints for **3** were calculated using phenix.elbow²⁷. The restraints for NADP⁺, **6**, and **8** are distributed with the CCP4 package and were used in their respective refinements. 3-Dimensional renderings of the structures were generated using the POVSCRIPT+ modification of

MOLSCRIPT^{28, 29}. The data collection and model refinement statistics are listed in Table 1. Coordinates have been deposited in the Protein Data Bank; the accession codes for each structure are also given in Table 1.

Results and Discussion

Structure determination of EcGhrA with NADP⁺ alone and in the presence of α -keto acid substrates. Since there were no crystal structures of the *E. coli* glyoxylate/hydroxypyruvate reductase available, we began by determining the structure of the EcGhrA·NADP⁺ complex. This form of the enzyme crystallized with two molecules per asymmetric unit arranged as a homodimer, as observed in the structures of homologous glyoxylate/hydroxypyruvate reductases^{30, 31}. GhrA homologs from other organisms are known to be active as homodimers^{32, 33}. Both the overall fold (Figure 1A) and active site architecture (Figure 1B) of EcGhrA are very similar to those of enzymes like *Sinorhizobium meliloti* GhrA (PDB ID 5UOG³⁰) and *Pyrococcus furiosus* GhrA (PDB ID

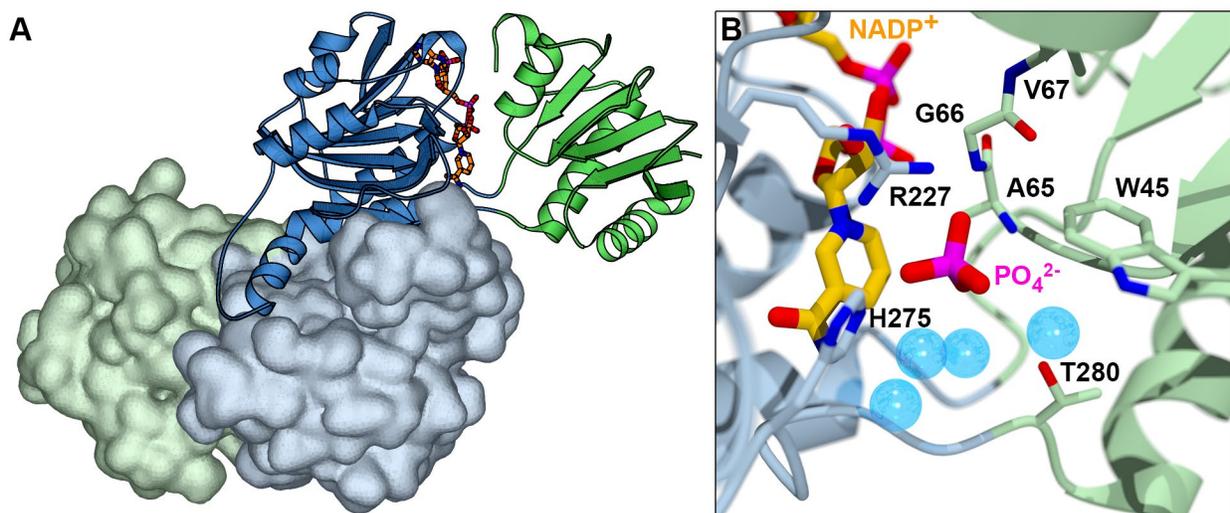


Figure 1. EcGhrA is active as a dimer of 35 kDa protomers as shown in (A). One protomer is shown in cartoon representation, and the other as a solvent-accessible surface. Each protomer is comprised of two domains: a nucleotide-binding domain (blue) and a substrate-binding domain (green). The active site of the EcGhrA·NADP⁺ binary complex (B) is organized exactly as observed in other GhrA homologs. The PO₄²⁻ anion bound in the active site is from the crystallization buffer.

5AOV³³). The main differences lie in the conformations of the dimerization loops (residues 111-136, EcGhrA numbering) and the C-terminal section beginning at residue 296.

The question driving this work—can we generate an EcGhrA variant that does not accept α -keto acids having large side chains (e.g. **3** and **8**) without abrogating activity with glyoxylate (**6**)—demanded that we determine how **6**, **3**, and **8** bind to the wild-type enzyme by determining structures of their ternary complexes. A challenge we encountered in this effort was that the crystallization conditions identified via high-throughput screening invariably contained high concentrations of phosphate, sulfate, or tartrate salts. We have observed that, at high concentrations, phosphate significantly inhibits EcGhrA activity (data not shown). Phosphate can be observed bound to R227 in the structure of the binary complex of EcGhrA with NADP⁺ (Figure 1B), and thus these ions compete with

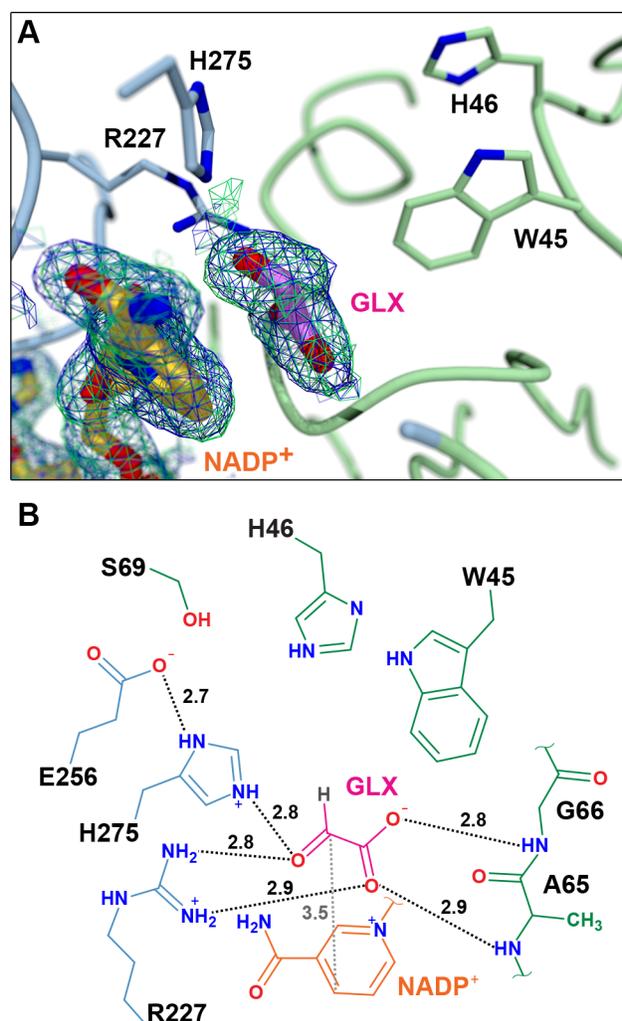


Figure 2. Electron density of NADP⁺ and **6** in the EcGhrA·NADP⁺·**6** ternary complex (A). The 2F_o - F_c electron density map contoured at 1.0 σ is shown as a blue mesh, and the 2F_o - F_c simulated annealing composite omit map, also contoured at 1.0 σ , is shown as a green mesh. The fact that the two maps are undistinguishable indicates that the electron density for the ligands is real and not the result of model bias. The schematic view of the ternary complex active site (B) shows hydrogen bonding interactions (black dotted lines) as well as the distance between the carbonyl carbon of **6** and C4 of the nicotinamide ring of NADP⁺ (grey dotted line). The distances in Ångströms are given next to each dotted line. Residues from the nucleotide-binding domain are colored blue, those from the substrate-binding domain are colored green, NADP⁺ is orange, and **6** is magenta.

substrates for binding at the active site. To solve this problem, we identified a new, phosphate-free crystallization condition. The new condition contained SO_4^- ions, but at a much lower concentration (0.2M) than the phosphate in the original condition (1.4M). When crystals grown from the phosphate-free condition were soaked with substrates as described in the Materials and Methods, the resulting crystal structures showed excellent electron density for the α -keto acid moieties of **6**, **3**, and **8** (Figures 2-4).

The EcGhrA·NADP⁺·**6** ternary complex structure was determined to elucidate the binding mode of the preferred substrate, glyoxylate (Figure 2). Glyoxylate binds such that the carbonyl and carboxylic acid groups participate in hydrogen bonding interactions with R227. The oxygen atoms of the carboxylic acid are also able to form hydrogen bonding interactions with the mainchain amides of A65 and G66. This binding mode orients the carbonyl carbon of glyoxylate 3.5 Å

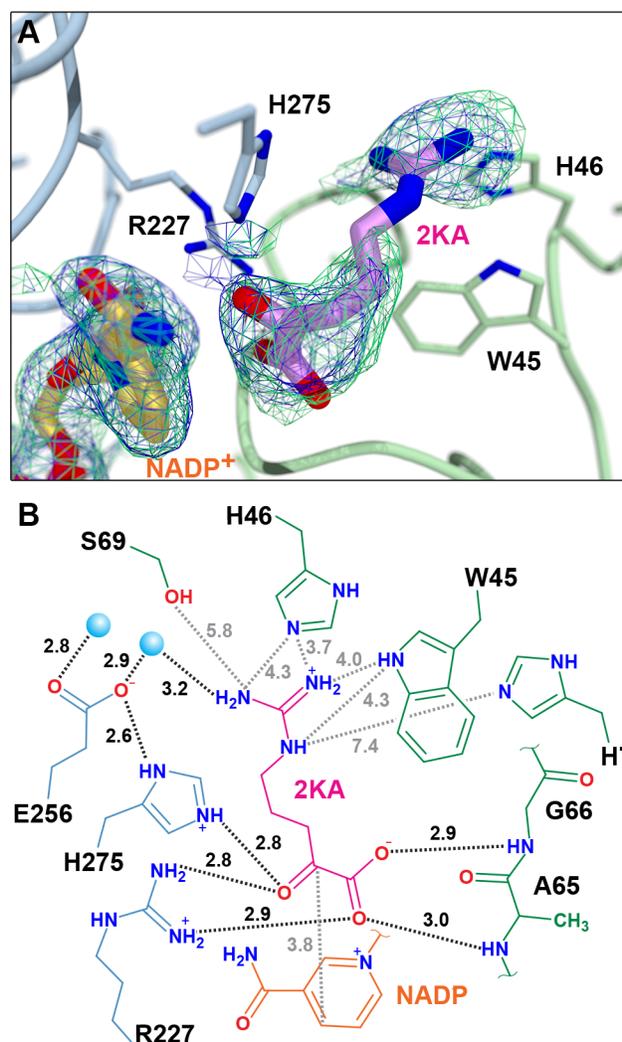


Figure 3. Electron density for the EcGhrA·NADP⁺·**3** ternary complex active site (A) using the same orientation as Figure 2. The 2F_o - F_c electron density map contoured at 1.0σ is shown as blue mesh. The 2F_o - F_c simulated annealing composite omit map (green mesh) is also contoured at 1.0σ. The schematic view (B) shows the hydrogen bonding interactions (black dotted lines) and their corresponding distances in Å. The lines do not represent non-covalent interactions, but simply indicate interesting distances. For example, the guanidinium group of **3** is relatively close (~4.0Å) to the indole N atom of W45. With a slight movement of the side chain, **3** might participate in cation-π interactions with this residue. As in Figure 2, residues from the nucleotide- and substrate-binding domains are blue and green, respectively. The NADP⁺ is orange and the substrate **3** is magenta. The blue spheres represent water molecules.

from the C4 position of the nicotinamide ring of the cosubstrate. This is exactly the binding mode observed in other GhrA homologs, such as the *P. furiosus* GhrA (PDB ID 5AOV,³³; Figure S2A). Unlike other GhrA homologs, EcGhrA does not appear to undergo a significant conformational change upon glyoxylate binding. NAD(H) binding does trigger a conformational change, since in

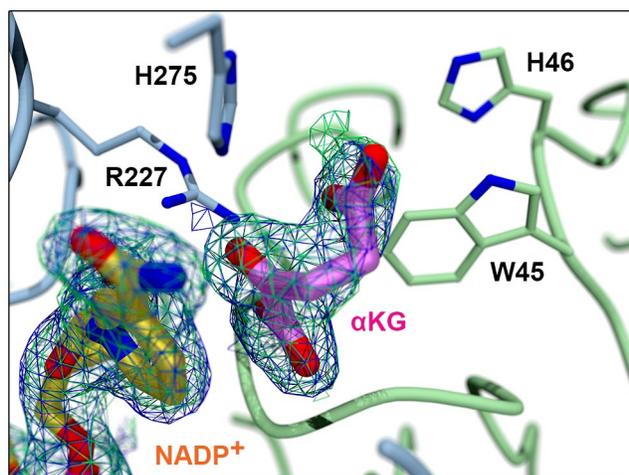


Figure 4. The EcGhrA·NADP⁺·**8** active site showing α -ketoglutarate bound adjacent to the cosubstrate. The $2F_o - F_c$ electron density map contoured at 1.0σ is shown as a blue mesh, and the $2F_o - F_c$ simulated annealing composite omit map, also contoured at 1.0σ , is shown as a green mesh. The side chain carboxylate group of **8** does not make any direct contacts with the enzyme.

the unliganded form of EcGhrA the substrate-binding domain (SBD) is completely disordered and only the nucleotide-binding domain (NBD) is visible (data not shown).

The crystal structure of the EcGhrA·NADP⁺·**3** ternary complex (Figure 3) was the most important with respect to informing our efforts to engineer EcGhrA to *not* accept **3** as a substrate. The structure shows that the keto acid moiety of **3** makes the same interactions with the enzyme as observed for

glyoxylate in the EcGhrA·NADP⁺·**6** ternary complex. Surprisingly, however, the guanidinium group of **3** makes no clear contacts with the enzyme (Figure 4B). While the position of the guanidinium group of **3** is defined in the electron density, there is little or no density for the methylene carbon atoms of the side chain. This suggests that there is considerable mobility in the side chain and binding is mediated primarily by the interactions of the α -keto acid moiety with the main chain amides of A65 and G66, and the side chains of R227 and H275 (Figure 3). It should be noted that even these ternary complex structures appear to be “open” when compared to those of other EcGhrA homologs with substrates bound (Figure S2). It is conceivable that the indole

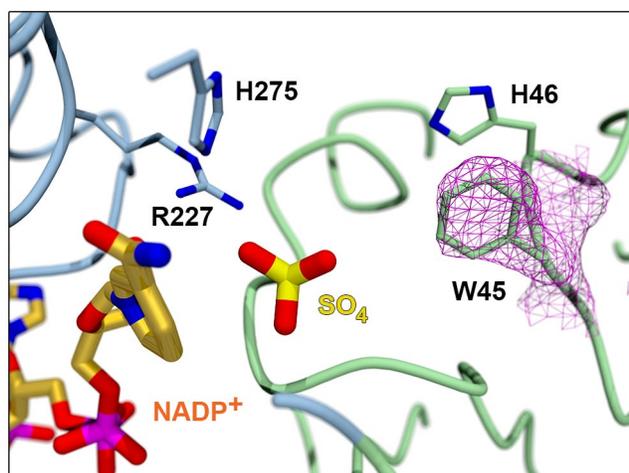


Figure 5. The EcGhrA^{W45F}·NADP⁺ active site showing the 2F_o - F_c electron density map around the W45F mutation (magenta mesh, contoured at 1.0σ). In the absence of substrate, the sulfate anion from the crystallization buffer binds near the cosubstrate.

nitrogen atom of W45 and/or the imidazole group of H46 could make strong contacts with the side chain of **3** if the substrate-binding domain were fully closed. The tryptophan would likely interact with the charged guanidinium group solely via cation- π interaction^{34, 35}, while the histidine could

accept a hydrogen bond from the protonated nitrogen atoms of the guanidinium group of **3**.

Based mostly on the possibility that W45 is

important for binding in a presumed “closed” conformation of EcGhrA, we set out to generate the W45F and W45F/H46S variants of the enzyme.

The structure of the ternary complex between EcGhrA, NADP⁺, and α -ketoglutarate (**8**) shows well-defined electron density for **8** near the nicotinamide ring of NADP⁺, in the same position occupied by glyoxylate in the structure of the EcGhrA·NADP⁺·**6** complex (Figure 4). The α -keto acid moiety makes the same interactions observed in the complexes with glyoxylate and 2-ketoarginine: the carboxylic acid forms hydrogen bonding interactions with the mainchain amides of A65 and G66, the guanidinium group of R227 bridges the carboxylate and ketone groups, and the imidazole group of H275 forms a hydrogen bonding interaction with the ketone of **8**. Interestingly, there is no clear indication from the structure why the activity with **8** is so much lower than the activity against glyoxylate. Likely this difference in reactivity is due to differences in the equilibria of the keto-enol tautomerizations in these two substrates.

Substrate specificity of EcGhrA. The kinetics of glyoxylate reduction catalyzed by wild-type

EcGhrA was studied first to establish a basis of comparison with other, more thoroughly characterized, GhrA homologs. EcGhrA exhibits a strong preference for NADPH over NADH (data not shown), so all kinetic analyses were done in the presence of saturating (200 μ M) NADPH. The k_{cat} value for the reaction of EcGhrA with glyoxylate (40.0 ± 23.0 ; Table 2, Figure S3) is similar to the values reported for the reaction of the *Homo sapiens* and *A. thaliana* homologs (27.0 and 25.5 s^{-1} , respectively)^{36, 37}. The K_{M} value for glyoxylate (2.2 ± 1.7 mM) was significantly different from those observed for the *H. sapiens* and *A. thaliana* enzymes (0.24 mM and 34 μ M, respectively)^{36, 37}. This difference may reflect differences in the structures and activities of these enzymes, or it may result from the acute substrate inhibition we observed (Figure S3; $K_{\text{i}}^6 = 1.5 \pm 1.1$ mM). The simple modified Michaelis-Menten equation (Equation 1) used to fit these data resulted in poor fits, as evidenced by the large standard uncertainties on the kinetic parameters. However, more nuanced models with additional parameters either failed to converge or resulted in absurdly large standard uncertainties.

Unlike glyoxylate (**6**), 2-ketoarginine (**3**) showed no sign of substrate inhibition. The turnover number for the reaction of EcGhrA with **3** (6.6 ± 0.4 s^{-1} ; Table 2, Figure S4) is only 6-fold lower than that for **6** and the K_{M} value for **3** is only 4-fold higher than that for **6**. The ratio of the pseudo-second order rate constants for **6** and **3** is 24 ± 4 , giving some quantitative idea of the relative preference of EcGhrA for these two substrates.

We used this ratio as a “selectivity factor” to compare the other substrates tested here (Scheme 2). Of compounds **7** – **12**, only oxaloacetate (**7**) and α -ketoglutarate (**8**) proved to be substrates for the enzyme. Oxaloacetate is nearly as efficient a substrate as **6**, and shows some evidence of substrate inhibition, though it is very weak ($K_{\text{i}}^7 = 52 \pm 15$ mM; Table 2, Figure S5). The turnover number for **7** (1.9 ± 0.1 s^{-1}) is approximately 20-fold lower than k_{cat}^6 . The K_{M} value for **7** (1.3 ± 0.2

mM) is roughly half of **6**. The selectivity factor for oxaloacetate is only 13 ± 3 , which underscores the similar catalytic efficiencies of oxaloacetate and glyoxylate. Conversely, α -ketoglutarate is a much poorer substrate, having a selectivity factor of 266 ± 50 (Table 2, Figure S6). With a K_M value only twice that for **6**, most of the difference between **8** and **6** comes in the turnover number. The value of k_{cat} ⁸ ($0.3 \pm 0.02 \text{ s}^{-1}$) is approximately 130-fold lower than the turnover number for glyoxylate. The K_M values for **8** and **3** are not significantly different, which is consistent with the close agreement between the positioning of **6** and **8** in the crystal structures.

The W45F variant of EcGhrA is more selective for glyoxylate. The W45F and H46S mutations were chosen as the most conservative changes that would still disrupt any potential interactions with substrate **3**. The W45F and W45F/H46S variants both expressed and purified similarly to the wild-type enzyme. There was no indication that mutation(s) had any effect on the stability of the enzyme. Steady state kinetic studies of both variants showed that the double mutant W45F/H46S did not behave differently from the single W45F variant. Therefore, only the W45F variant will be discussed here. EcGhrA^{W45F} exhibited a roughly 10-fold decrease in k_{cat}/K_M for **3**, while the k_{cat}/K_M for **6** turned out to be statistically within the range of the values for the wild-type (Table 2). Thus, the W45F mutation somewhat decreased the catalytic efficiency of the reaction with **3**, while leaving the reaction with **6** essentially untouched. The selectivity factor for **6** with respect to **3** increased 5-fold for the mutant (119 vs 24 for the wild-type). It is also important to mention that the K_M values for both **6** and **3** are likely overestimates, since **6** exists predominantly in a hydrated gem-diol form³⁸, while **3** is in equilibrium with the cyclic pyrrolidine-1-amidino-2-hydroxy-2-carboxylic acid³⁹.

The crystal structure of the single mutant showed no conformational changes of the mutated residues, or the catalytic residues R227 and H275 (Figure 5), suggesting that the substrate binding

modes remain unchanged. The single W45 to Phe mutation did not disrupt the native EcGhrA reaction but did significantly reduce the reactivity with **3**. The resulting EcGhrA^{W45F} variant is thus more selective for glyoxylate and is a suitable coupling enzyme for steady state assays.

Table 1: Crystallographic data collection and refinement statistics.

	EcGhrA·NAD P ⁺	EcGhrA·NAD P ⁺ ·3	EcGhrA·NAD P ⁺ ·8	EcGhrA·NAD P ⁺ ·6	EcGhrA ^{W45F} ·N ADP ⁺
PDB Entry	7JQH	6P35	7JQI	6OVL	7JQJ
<i>Data collection</i>					
Resolution (Å) (last shell)^a	46.03 – 1.999 (2.07 -1.999)	45.78 – 2.5 (2.589 – 2.5)	45.87 – 2.399 (2.485 – 2.399)	36.73 – 2.1 (2.175 -2.1)	35.24 – 2.204 (2.283 – 2.204)
Space group	P 65 2 2	P 65 2 2	P 65 2 2	P 65 2 2	P 65 2 2
a, b, c (Å)	158.768, 158.768, 97.699	158.578, 158.578, 98.285	158.887, 158.887, 96.091	158.82, 158.82, 96.71	159.177, 159.77, 95.709
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
R_{merge}^a	0.1228 (0.7386)	0.1459 (0.8655)	0.1079 (0.3597)	0.09166 (0.7968)	0.116 (0.5544)
R_{meas}^a	0.1257 (0.7557)	0.1495 (0.8901)	0.1148 (0.382)	0.09453 (0.8394)	0.1201 (0.5903)
R_{pim}^a	0.02673 (0.1594)	0.03261 (0.2057)	0.03891 (0.1277)	0.02256 (0.2548)	0.03031 (0.1956)
CC_{1/2}^a	0.998 (0.942)	0.998 (0.903)	0.995 (0.948)	0.999 (0.897)	0.998 (0.853)
No. of unique reflections^a	49390 (4836)	25704 (2516)	28341 (2727)	42222 (3925)	36363 (2919)
Completeness (%)^a	99.95 (99.77)	99.83 (99.88)	99.44 (97.77)	99.09 (94.62)	97.28 (81.59)
Multiplicity^a	22.1 (22.3)	20.5 (18.4)	8.5 (8.7)	16.2 (9.6)	15.1 (8.3)
$\langle I/\sigma(I) \rangle^a$	26.24 (4.65)	24.37 (3.57)	22.70 (6.32)	29.83 (3.51)	16.11 (2.33)
<i>Model Refinement</i>					
Reflections used in refinement^a	49371 (4836)	25669 (2516)	28324 (2723)	41937 (3925)	35509 (2920)
Reflections used for R_{free}^a	2000 (196)	1992 (196)	1403 (157)	2121 (214)	1763 (148)
R_{cryst} (R_{free})^a	0.1735 (0.2095)	0.2057 (0.2585)	0.1965 (0.2135)	0.1835 (0.2451)	0.1950 (0.2479)
Wilson B-factor (Å²)	24.25	36.50	34.92	27.74	32.44
Average B factor (Å²)	28.46	41.52	39.56	32.67	37.77
Protein atoms	27.34	41.44	39.43	31.67	37.54
Ligand atoms	23.00	49.19	33.27	29.78	26.23

Solvent	38.75	36.92	42.86	42.34	43.15
Root-mean-square (RMS) deviations					
Bond lengths (Å)	0.014	0.012	0.011	0.004	0.005
Bond angles (°)	1.15	1.33	1.13	0.99	1.01
Coordinate error (Å)^b	0.14	0.26	0.20	0.19	0.20
Ramachandran statistics					
Favored/allowed/outliers (%)	97.43/2.25/0.3 2	97.74/1.61/0.6 5	96.77/2.90/0.3 2	97.43/2.25/0.3 2	97.06/2.94/0.0 0
Rotamer outliers (%)	0.00	2.64	0.00	0.38	0.00
Clashscore	1.00	2.94	1.60	0.20	1.02
TLS groups	7	3	10	1	4

^a Values in parentheses apply to the high-resolution shell indicated in the resolution row.

^b $R = \sum (||F_{obs}| - scale * |F_{calc}||) / \sum |F_{obs}|$.

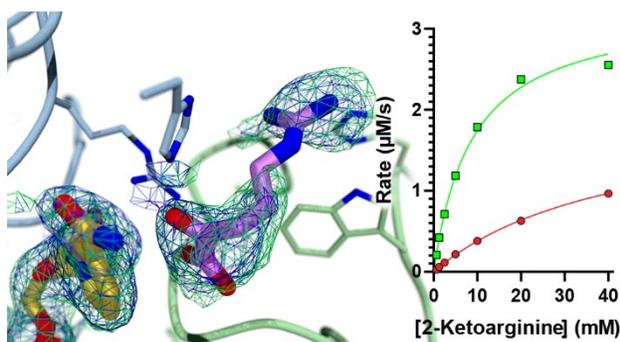
^c Isotropic-equivalent B factors, including contribution from TLS refinement.

Table 2. Steady state parameters of EcGhrA and EcGhrA^{W45F} with **6**, **3**, **7**, and **8**.

	EcGhrA	EcGhrA^{W45F}
k_{cat}^6 (s ⁻¹)	40.0 ± 23.0	120.0 ± 63.0
K_M^6 (mM)	2.2 ± 1.7	10.8 ± 6.1
k_{cat}^6/K_M^6 (M ⁻¹ s ⁻¹)	(2.0 ± 1.8) × 10 ⁴	(1.1 ± 0.9) × 10 ⁴
K_I^6 (mM)	1.5 ± 1.1	4.3 ± 2.9
k_{cat}^3 (s ⁻¹)	6.6 ± 0.4	4.0 ± 0.7
K_M^3 (mM)	8.8 ± 0.8	43.2 ± 8.4
k_{cat}^3/K_M^2 (M ⁻¹ s ⁻¹)	(8.3 ± 1.0) × 10 ²	92.5 ± 24.0
K_I^3 (mM)	NA	NA
Selectivity factor ^a	24 ± 4	119 ± 32
k_{cat}^7 (s ⁻¹)	1.9 ± 0.1	1.8 ± 0.04
K_M^7 (mM)	1.3 ± 0.2	6.3 ± 0.2
k_{cat}^7/K_M^7 (M ⁻¹ s ⁻¹)	(1.5 ± 0.3) × 10 ³	290 ± 17
K_I^7 (mM)	52 ± 15	NA
Selectivity factor ^a	13 ± 3	38 ± 4
k_{cat}^8 (s ⁻¹)	0.30 ± 0.02	0.26 ± 0.01
K_M^8 (mM)	4.0 ± 0.4	11.5 ± 0.8
k_{cat}^8/K_M^8 (M ⁻¹ s ⁻¹)	75.2 ± 12.5	22.6 ± 2.3
K_I^8 (mM)	NA	NA
Selectivity factor ^a	266 ± 50	487 ± 64

^a The selectivity factor is taken as the ratio of the pseudo-second order rate constant for **6** to those of the other substrates tested here.

TABLE OF CONTENTS USE ONLY



REFERENCES

- [1] Singh, M. P., Petersen, P. J., Weiss, W. J., Janso, J. E., Luckman, S. W., Lenoy, E. B., Bradford, P. A., Testa, R. T., and Greenstein, M. (2003) Mannopectimycins, new cyclic glycopeptide antibiotics produced by *Streptomyces hygroscopicus* LL-AC98: antibacterial and mechanistic activities, *Antimicrob. Agents Chemother.* 47, 62-69.
- [2] He, H., Williamson, R. T., Shen, B., Graziani, E. I., Yang, H. Y., Sakya, S. M., Petersen, P. J., and Carter, G. T. (2002) Mannopectimycins, novel antibacterial glycopeptides from *Streptomyces hygroscopicus*, LL-AC98, *J Am Chem Soc* 124, 9729-9736.
- [3] Han, L., Vuksanovic, N., Oehm, S. A., Fenske, T. G., Schwabacher, A. W., and Silvaggi, N. R. (2018) *Streptomyces wadayamensis* MppP is a PLP-Dependent Oxidase, Not an Oxygenase, *Biochemistry* 57, 3252-3264.
- [4] Han, L., Schwabacher, A. W., Moran, G. R., and Silvaggi, N. R. (2015) *Streptomyces wadayamensis* MppP Is a Pyridoxal 5'-Phosphate-Dependent L-Arginine alpha-Deaminase, gamma-Hydroxylase in the Enduracididine Biosynthetic Pathway, *Biochemistry* 54, 7029-7040.
- [5] Burroughs, A. M., Hoppe, R. W., Goebel, N. C., Sayyed, B. H., Voegtline, T. J., Schwabacher, A. W., Zabriskie, T. M., and Silvaggi, N. R. (2013) Structural and functional characterization of MppR, an enduracididine biosynthetic enzyme from *streptomyces hygroscopicus*: functional diversity in the acetoacetate decarboxylase-like superfamily, *Biochemistry* 52, 4492-4506.

- [6] el-Mansi, E. M., MacKintosh, C., Duncan, K., Holms, W. H., and Nimmo, H. G. (1987) Molecular cloning and over-expression of the glyoxylate bypass operon from *Escherichia coli* ML308, *Biochem J* 242, 661-665.
- [7] Pineda, M., Piedras, P., and Cardenas, J. (1994) A continuous spectrophotometric assay for ureidoglycolase activity with lactate dehydrogenase or glyoxylate reductase as coupling enzyme, *Anal Biochem* 222, 450-455.
- [8] Matelska, D., Shabalin, I. G., Jablonska, J., Domagalski, M. J., Kutner, J., Ginalski, K., and Minor, W. (2018) Classification, substrate specificity and structural features of D-2-hydroxyacid dehydrogenases: 2HADH knowledgebase, *BMC Evol Biol* 18, 199.
- [9] Nunez, M. F., Pellicer, M. T., Badia, J., Aguilar, J., and Baldoma, L. (2001) Biochemical characterization of the 2-ketoacid reductases encoded by *ycdW* and *yiaE* genes in *Escherichia coli*, *Biochem J* 354, 707-715.
- [10] Yokota, A., Haga, S., and Kitaoka, S. (1985) Purification and some properties of glyoxylate reductase (NADP⁺) and its functional location in mitochondria in *Euglena gracilis* z, *Biochem J* 227, 211-216.
- [11] Tochikura, T., Fukuda, H., and Moriguchi, M. (1979) Purification and properties of glyoxylate reductase I from baker's yeast, *J Biochem* 86, 105-110.
- [12] Allan, W. L., Simpson, J. P., Clark, S. M., and Shelp, B. J. (2008) Gamma-hydroxybutyrate accumulation in *Arabidopsis* and tobacco plants is a general response to abiotic stress: putative regulation by redox balance and glyoxylate reductase isoforms, *J Exp Bot* 59, 2555-2564.

- [13] Fukuda, H., Moriguchi, M., Kimura, A., and Tochikura, T. (1981) Purification and Properties of Glyoxylate Reductase from *Neurospora crassa*, *Agricultural and Biological Chemistry* 45, 1153-1158.
- [14] Shinoda, T., Arai, K., and Taguchi, H. (2007) A highly specific glyoxylate reductase derived from a formate dehydrogenase, *Biochem Biophys Res Commun* 355, 782-787.
- [15] Zheng, Z., Sheng, B., Gao, C., Zhang, H., Qin, T., Ma, C., and Xu, P. (2013) Highly stereoselective biosynthesis of (R)-alpha-hydroxy carboxylic acids through rationally re-designed mutation of D-lactate dehydrogenase, *Sci Rep* 3, 3401.
- [16] Wei, L., Cai, X., Qi, Z., Rong, L., Cheng, B., and Fan, J. (2012) In vivo and in vitro characterization of TEV protease mutants, *Protein Expr Purif* 83, 157-163.
- [17] Meister, A. (1952) Enzymatic preparation of alpha-keto acids, *J Biol Chem* 197, 309-317.
- [18] Stalon, V., and Mercenier, A. (1984) L-arginine utilization by *Pseudomonas* species, *J Gen Microbiol* 130, 69-76.
- [19] Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol* 276, 307-326.
- [20] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J Appl Crystallogr* 40, 658-674.
- [21] Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr D Biol Crystallogr* 67, 235-242.

- [22] Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R. W., Moriarty, N. W., Langan, P., and Adams, P. D. (2010) Joint X-ray and neutron refinement with phenix.refine, *Acta Crystallogr D Biol Crystallogr* 66, 1153-1163.
- [23] Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr* 66, 213-221.
- [24] Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132.
- [25] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr D Biol Crystallogr* 66, 486-501.
- [26] Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr* 66, 12-21.
- [27] Moriarty, N. W., Grosse-Kunstleve, R. W., and Adams, P. D. (2009) electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation, *Acta Crystallogr D Biol Crystallogr* 65, 1074-1080.
- [28] T. D. Fenn, D. R. G. A. P. (2004) POVScript+: a program for model and data visualization using persistence of vision ray-tracing, *Journal of Applied Crystallography* 36, 944-947.

- [29] Kraulis, P. J., and Iucr. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *Journal of Applied Crystallography* 24, 946-950.
- [30] Kutner, J., Shabalin, I. G., Matelska, D., Handing, K. B., Gasiorowska, O., Sroka, P., Gorna, M. W., Ginalski, K., Wozniak, K., and Minor, W. (2018) Structural, Biochemical, and Evolutionary Characterizations of Glyoxylate/Hydroxypyruvate Reductases Show Their Division into Two Distinct Subfamilies, *Biochemistry* 57, 963-977.
- [31] Booth, M. P., Conners, R., Rumsby, G., and Brady, R. L. (2006) Structural basis of substrate specificity in human glyoxylate reductase/hydroxypyruvate reductase, *J Mol Biol* 360, 178-189.
- [32] Ogino, H., Nakayama, H., China, H., Kawata, T., Doukyu, N., and Yasuda, M. (2008) Characterization of recombinant glyoxylate reductase from thermophile *Thermus thermophilus* HB27, *Biotechnol Prog* 24, 321-325.
- [33] Lassalle, L., Engilberge, S., Madern, D., Vauclare, P., Franzetti, B., and Girard, E. (2016) New insights into the mechanism of substrates trafficking in Glyoxylate/Hydroxypyruvate reductases, *Sci Rep* 6, 20629.
- [34] Gallivan, J. P., and Dougherty, D. A. (1999) Cation-pi interactions in structural biology, *Proc Natl Acad Sci U S A* 96, 9459-9464.
- [35] Dougherty, D. A. (2007) Cation-pi interactions involving aromatic amino acids, *J Nutr* 137, 1504S-1508S; discussion 1516S-1517S.
- [36] Mdluli, K., Booth, M. P., Brady, R. L., and Rumsby, G. (2005) A preliminary account of the properties of recombinant human Glyoxylate reductase (GRHPR), LDHA and LDHB

with glyoxylate, and their potential roles in its metabolism, *Biochim Biophys Acta* 1753, 209-216.

- [37] Simpson, J. P., Di Leo, R., Dhanoa, P. K., Allan, W. L., Makhmoudova, A., Clark, S. M., Hoover, G. J., Mullen, R. T., and Shelp, B. J. (2008) Identification and characterization of a plastid-localized Arabidopsis glyoxylate reductase isoform: comparison with a cytosolic isoform and implications for cellular redox homeostasis and aldehyde detoxification, *J Exp Bot* 59, 2545-2554.
- [38] Duncan, R. J., and Tipton, K. F. (1969) The oxidation and reductn of glyoxylate by lactic dehydrogenase, *Eur J Biochem* 11, 58-61.
- [39] Cooper, A. J., and Meister, A. (1978) Cyclic forms of the alpha-keto acid analogs of arginine, citrulline, homoarginine, and homocitrulline, *J. Biol. Chem.* 253, 5407-5410.