BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

# The Structure and Function of a Microbial Allantoin Racemase Reveal the Origin and Conservation of a Catalytic Mechanism

Laura Cendron,[†,∥] Ileana Ramazzina,[‡,⊥] Vincenzo Puggioni,[‡] Eleonora Maccacaro,[‡] Anastasia Liuzzi,[‡] Andrea Secchi,[§] Giuseppe Zanotti,[†] and Riccardo Percudani*[,‡]

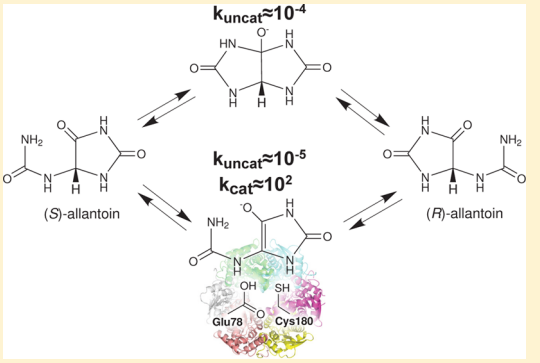
[†]Department of Biomedical Sciences, University of Padova, Padova, Italy

[‡]Department of Life Sciences, University of Parma, Parma, Italy

[§]Department of Chemistry, University of Parma, Parma, Italy

Supporting Information

**ABSTRACT:** The *S* enantiomer of allantoin is an intermediate of purine degradation in several organisms and the final product of uricolysis in nonhominoid mammals. Bioinformatics indicated that proteins of the Asp/Glu racemase superfamily could be responsible for the allantoin racemase (AllR) activity originally described in *Pseudomonas* species. In these proteins, a cysteine of the catalytic dyad is substituted with glycine, yet the recombinant enzyme displayed racemization activity with a similar efficiency ($k_{cat}/K_M \approx 5 \times 10^4$ $M^{-1}$ $s^{-1}$) for the *R* and *S* enantiomers of allantoin. The protein crystal structure identified a glutamate residue located three residues downstream (E78) that can functionally replace the missing cysteine; the catalytic role of E78 was confirmed by site-directed mutagenesis. Allantoin can undergo racemization through formation of a bicyclic intermediate (faster) or proton exchange at the chiral center (slower). By monitoring the two alternative mechanisms by $^{13}C$ and $^{1}H$ nuclear magnetic resonance, we found that the velocity of the faster reaction is unaffected by the enzyme, whereas the velocity of the slower reaction is increased by 7 orders of magnitude. Protein phylogenies trace the origin of the racemization mechanism in enzymes acting on glutamate, a substrate for which proton exchange is the only viable reaction mechanism. This mechanism was inherited by allantoin racemase through divergent evolution and conserved in spite of the substitution of catalytic residues.




A single enantiomer of allantoin is produced in nature through uricolysis, a three-step purine degradation pathway whose last reaction is the enantioselective decarboxylation of 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline (OHCU) to give (*S*)-allantoin.[1−4] With few exceptions,[5] the enzymes for allantoin breakdown (allantoinase), though encoded by two different gene families, specifically catalyze hydrolysis of the *S* enantiomer.[6,7] However, racemic allantoin is produced by the chemical oxidation of urate or its incomplete enzymatic conversion catalyzed by urate oxidase.[8,9] In addition, if allantoin is not hydrolyzed but rather eliminated in the environment, the *R* enantiomer is produced by spontaneous racemization, which occurs at a rate constant of $\sim 2 \times 10^{-5}$ at neutral pH.[10] Remarkably, two separate mechanisms exist for allantoin racemization: formation of a bicyclic intermediate that can decompose to form either enantiomer of allantoin and proton exchange at the chiral center with solvent. Both mechanisms contribute to spontaneous racemization, but the pathway through the bicyclic intermediate is ~10 times faster than the other.[10] Allantoin can be utilized by various organisms as a nutrient source.[11] Only half of the racemic mixture can be readily utilized through stereospecific allantoinase, unless an enzyme is present that is able to catalyze the interconversion of allantoin enantiomers.

Enzymes with allantoin racemase (EC 5.1.99.3) activity have a potential use in the determination of the output of urate oxidation. (*S*)-Allantoin can be quantified in serum by an enzymatic method based on the use of stereospecific allantoinase.[12] The use of allantoin racemase would allowe the quantification of both enantiomers and the distinction between enantioselective and non-enantioselective urate degradation. This can be useful for monitoring uricolytic activities in biological fluids after administration of uricolytic drugs in the enzymatic therapy of severe hyperuricemia caused by tumor lysis syndrome and Lesch-Nyhan disease.[13,14]

In early studies, isolation of the (*R*)-allantoin enantiomer through the use of stereospecific allantoinase[15] allowed several *Pseudomonas* species to be tested for growth on this substrate; some species were found to utilize this substrate as a nitrogen source because of the presence of a racemase enzyme.[5] More recently, the *Dcg1p* gene of *Saccharomyces cerevisiae* and the *hpxA* gene of *Klebsiella* spp. were predicted to encode allantoin racemase (AllR) for their proximity with other genes involved

**Received:** September 1, 2016
**Revised:** October 30, 2016
**Published:** October 31, 2016

ACS Publications © XXXX American Chemical Society



DOI: 10.1021/acs.biochem.6b00881
*Biochemistry* XXXX, XXX, XXX−XXX

in purine degradation and sequence similarity with other racemases.[16,17]

The demonstration of the AllR function was provided by the functional and structural characterization of the *Klebsiella pneumoniae* enzyme.[18] The *Kp*AllR protein is able to catalyze allantoin racemization with similar $k_{cat}$ and $K_M$ values for both enantiomers. The structure of the protein in complex with allantoin identifies two cysteine residues (C79 and C184) situated on opposite sides of the substrate chiral center in a position consistent with the two-base mechanism typical of the Asp/Glu racemase superfamily. Replacement of either of the two cysteines with serine leads to the loss of enzymatic activity.[18] The *Kp*AllR mechanism is confirmed by quantum mechanics/molecular mechanics (QM/MM) calculations of the enzyme-catalyzed reaction, suggesting that allantoin stereoinversion proceeds through a stepwise mechanism involving a transient unprotonated intermediate.[19]

Here we present evidence, obtained through bioinformatics analysis, of the identification at the molecular level of the allantoin racemase enzyme (*Pf*AllR) originally described in *Pseudomonas* species. This evidence includes sequence similarity with the characterized *Kp*AllR protein. Interestingly, however, the candidate *Pf*AllR and orthologous proteins show the substitution of the first cysteine of the catalytic dyad with glycine. In spite of this nonconservative substitution, nuclear magnetic resonance (NMR) analysis demonstrated that the *Pf*AllR enzyme can catalyze proton exchange at the allantoin chiral center, while not affecting the velocity of formation of the bicyclic intermediate, an alternative faster mechanism of allantoin racemization. The atomic structure of the *Pf*AllR enzyme provided a rationale for the enzyme activity, identifying a glutamate residue that can functionally replace the substituted cysteine in catalysis. The evolution of allantoin racemase within the Asp/Glu racemase superfamily illustrates a clear example of the importance of the enzyme ancestry for the choice and conservation of a particular catalytic mechanism.

## ■ EXPERIMENTAL PROCEDURES

**Bioinformatics.** The analysis of genotype/phenotype correlation was conducted with the OrthologueDB web server (www.pathogenomics.sfu.ca/orthologuedb)[20] by selecting completely sequenced organisms that were reported to possess or not possess AllR activity.[5] The complete set of orthologous genes was "filtered" according to the reported phenotype to obtain a list of orthologs present in AllR+ and absent in AllR− organisms. The presence of the identified genes in three genomes not included in Orthologuedb (*Pseudomonas tolaasii*, *Pseudomonas alcaligenes*, and *Pseudomonas pseudoalcaligenes*) was determined through reciprocal best hit blast searches. Comparative analysis of the gene context was performed with the MicrobesOnline web server (http://www.microbesonline.org).[21] Sequence alignments and phylogeneic trees (based on neighbor joining of corrected genetic distances) were generated using Clustalx.[22] Sequence logos for different groups of the Asp/Glu superfamiliy were generated using the WebLogo server (http://weblogo.berkeley.edu).[23]

**Protein Expression and Purification.** DNA encoding the candidate AllR was amplified from a *Pseudomonas fluorescens* type strain (DSM 50090). Genomic DNA was prepared using described procedures,[24] and the region encoding putative AllR protein was amplified by polymerase chain reaction using Go Taq DNA polymerase (Promega) and two sequence specific primers (Table S2). The amplification product was inserted into the pGEM vector (Promega) to generate intermediate vector pGEM-AllR. The restriction fragment obtained from NdeI and BamHI digestion was then ligated into expression vector pET11b (Novagen). The resulting plasmid was verified by sequencing and electroporated into *Escherichia coli* BL21 codon plus cells (DE3). The expression of AllR was induced at an optical density at 600 nm of 0.6 with 1 mM isopropyl β-D-1-thiogalactopyranoside; after 5 h at 28 °C, the cells were resuspended in 180 mL of lysis buffer [50 mM sodium phosphate, 0.3 M NaCl, 1 mM β-mercaptoethanol, 10% glycerol, 1 μM pepstatin, 1 μM leupeptin, and 100 μM phenylmethanesulfonyl fluoride (pH 8)] and lysed by sonication. The supernatant obtained after centrifugation of the crude extract was concentrated by ultrafiltration in an Amicon cell (YM-10 membrane, Millipore). AllR was purified to homogeneity, as assessed by 16% sodium dodecyl sulfate−polyacrylamide gel electrophoresis analysis, by anion exchange (Q Sepharose, Pharmacia) and gel filtration (AcA54, Pharmacia) chromatography, with a final yield of approximately ∼6 mg/L of cell culture. The anion exchange column (10 cm × 1.5 cm) was equilibrated in 0.1 M KP (pH 7.6) and eluted with a linear gradient of NaCl (0 to 0.5 M in 60 mL of column buffer). Fractions corresponding to ∼0.3 M NaCl were concentrated by ultrafiltration in an Amicon cell (YM-10 membrane, Millipore), applied to the AcA54 column (120 cm × 1.5 cm), and eluted with 0.3 M NaCl in 0.1 M KP (pH 7.6). $\varepsilon_{280}$ (extinction coefficient) of AllR was estimated to be 15930 $M^{-1}$ $cm^{-1}$, based on the amino acid sequence. The molecular weight of purified AllR was estimated by gel filtration on a Sephadex G-100 column (Pharmacia) (55 cm × 1.5 cm), using 0.1 M potassium phosphate (pH 7.6) as buffer for equilibration and elution.

**Crystallization and Data Collection.** Initial crystallization screens were performed at 20 °C by a vapor diffusion technique using an automated sitting-drop setup (Oryx-8, Douglas Instruments). The most promising conditions were reproduced and optimized by the hanging-drop method. The best crystals grew under two conditions of the Structure screen I and II HT-96 kit (Molecular Dimension Ltd.): SSI n.38 [0.1 M Tris (pH 8.5) and 8% (w/v) PEG 8000] and SSI n.23 [0.2 M $CaCl_2$, 0.1 M HEPES (pH 7.5), and 28% (w/v) PEG 400], corresponding to two different crystal forms that were grown. Cubic crystals ($a = b = c = 109.72$ Å), grown under condition SSI n.38, belong to the $P2_13$ space group and contain two monomers per asymmetric unit ($V_M$ = 2.21 $Å^3$ $Da^{-1}$, solvent content of 44%). Orthorhombic crystals ($P2_12_12_1$, $a$ = 60.22 Å, $b$ = 142.32 Å, $c$ = 146.02 Å), grown under condition SSI n.23, contain an entire hexamer in the asymmetric unit, corresponding to a $V_M$ of 2.08 $Å^3$ $Da^{-1}$ and a solvent content of ∼41%. Data for the cubic and orthorhombic crystal forms were measured at beamlines ID23-1 and ID14-1, respectively, of the European Synchrotron Radiation Facility (ESRF, Grenoble, France). Both crystal forms were cryoprotected by soaking for a few seconds in the corresponding crystallization condition supplemented with 20% glycerol, prior to being flash-frozen and stored in liquid nitrogen. The best crystals diffracted to maximal resolutions of 2.10 and 2.15 Å for the orthorhombic and cubic crystal forms, respectively.

**Structure Determination and Refinement.** All data sets were indexed and integrated with Mosflm software[25] and merged and scaled with Scala,[26] contained in the CCP4 crystallographic package.[27] Both structures were determined by molecular replacement using the Phaser software.[28] The model


DOI: 10.1021/acs.biochem.6b00881
Biochemistry XXXX, XXX, XXX−XXX

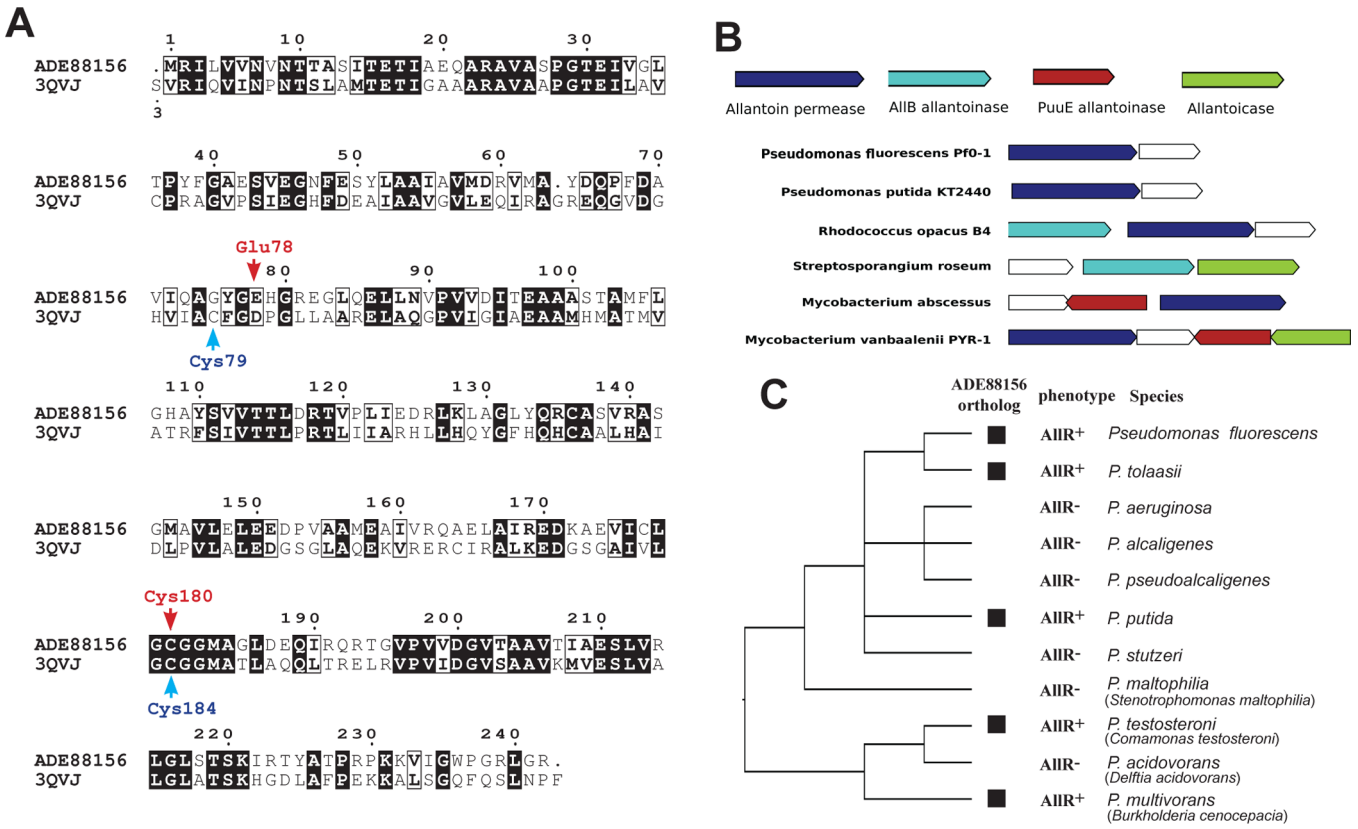



**Figure 1.** Bioinformatics evidence of the identification of *P. fluorescens* AllR. (A) Alignment of *K. pneumoniae* AllR (Protein Data Bank entry 3QVJ) with the homologous *P. flourescens* protein (GenBank accession number ADE88156). Known residues involved in the two-base catalysis are denoted with blue arrows; functionally equivalent residues identified in the *Pf*AllR structure (see Figure 3D) are denoted with red arrows. (B) Examples of genetic associations of the putative *Pf*AllR and similar genes (white boxes) with genes involved in purine catabolism (colored boxes). (C) Genotype/phenotype correlation between the presence of *Pf*AllR orthologs in complete genomes and the AllR phenotype. The tree is in accordance with the NCBI taxonomy.

was displayed and manually adjusted with graphic Coot software.[29] Refinement was conducted using Refmac[30] and Phenix[31] packages. The two structures were refined to final *R* factors of 0.205 ($R_{free}$ = 0.264) and 0.188 ($R_{free}$ = 0.233) for the cubic and orthorhombic crystal forms, respectively. Geometrical parameters of the models, checked with Procheck,[32] are as expected or better for this resolution. Data collection and refinement statistics are listed in Table 3.

**Synthesis of Allantoin Enantiomers.** (*S*)-Allantoin was synthesized enzymatically from urate. A typical reaction mixture [10 mL, 100 mM KP (pH 7.6)] contained 5 mg/mL urate, 10 mg of Uox from *Candida* (Sigma), 2.5 mg of catalase, and recombinant Urah (0.1 mg) and Urad (1 mg) from zebrafish.[2,33] The mix was incubated while being stirred at room temperature; the complete conversion of uric acid to allantoin was monitored by recording spectra between 200 and 340 nm. (*R*)-Allantoin was synthesized enzymatically starting from racemic allantoin. A typical reaction mixture [10 mL, 100 mM KP (pH 7.6)] contained 10 mg/mL allantoin and recombinant allantoinase (0.2 mg) from *P. fluorescens*.[6] Proteins were removed from the solution by Amicon (10 kDa cutoff). The prepared enantiomers were utilized immediately in enzymatic assays or briefly stored at −80 °C for the determination of the reaction velocity at different substrate concentrations. The (*R*)-allantoin preparation contained allantoic acid; however, the AllR reaction was reported not to be inhibited by this compound.[18]

**Site-Directed Mutagenesis.** Site-directed mutants E78Q and E78D were obtained by site-directed mutagenesis using a mix (50 μL) containing 50 ng of the *Pf*AllR expression vector, 2.5 units of PFU turbo (Stratagene, La Jolla, CA), 0.3 mM dNtp, 2% DMSO, and 126 ng of each mutagenic primers (Table S2). Reactions were performed as follows: initial denaturation at 95 °C (30 s) followed by 16 cycles of 95 °C for 30 s, 55 °C for 60 s, and 68 °C for 7 min, followed by 5 min at 68 °C. For each mutation, the product of the reaction was treated with DpnI (New England Biolabs) to digest the parental DNA template. The products of each digestion were used to transform *E. coli* XL1 Blue cells. Single clones were sequenced to confirm the presence of the desired mutation.

**CD Spectroscopy.** Circular dichroism measurements were taken at 25 °C in a 10 mm light path length cuvette with a Jasco J715 spectropolarimeter. AllR activity was determined using (*S*)-allantoin or (*R*)-allantoin as the substrate in the presence of AllR (9 μg) in 1 mL of 0.1 M potassium phosphate (pH 7.6). CD spectra were recorded every minute in the range of 200−300 nm for 10 min. To evaluate the kinetic parameters of wild-type and mutants enzymes, the ellipticity of allantoin enantiomers[6] was monitored at a fixed wavelength of 220 nm (molar ellipticity of 2.1 × $10^3$ deg $M^{-1}$ $m^{-1}$) for concentrations of ≤0.6 mM and of 240 nm (molar ellipticity of 0.2 × $10^3$ deg $M^{-1}$ $m^{-1}$) for higher concentrations. The reactions were initiated by the addition of AllR (1.8 μg) to 1 mL of 0.1 M potassium phosphate (pH 7.6). The reaction rate was obtained


DOI: 10.1021/acs.biochem.6b00881
Biochemistry XXXX, XXX, XXX−XXX

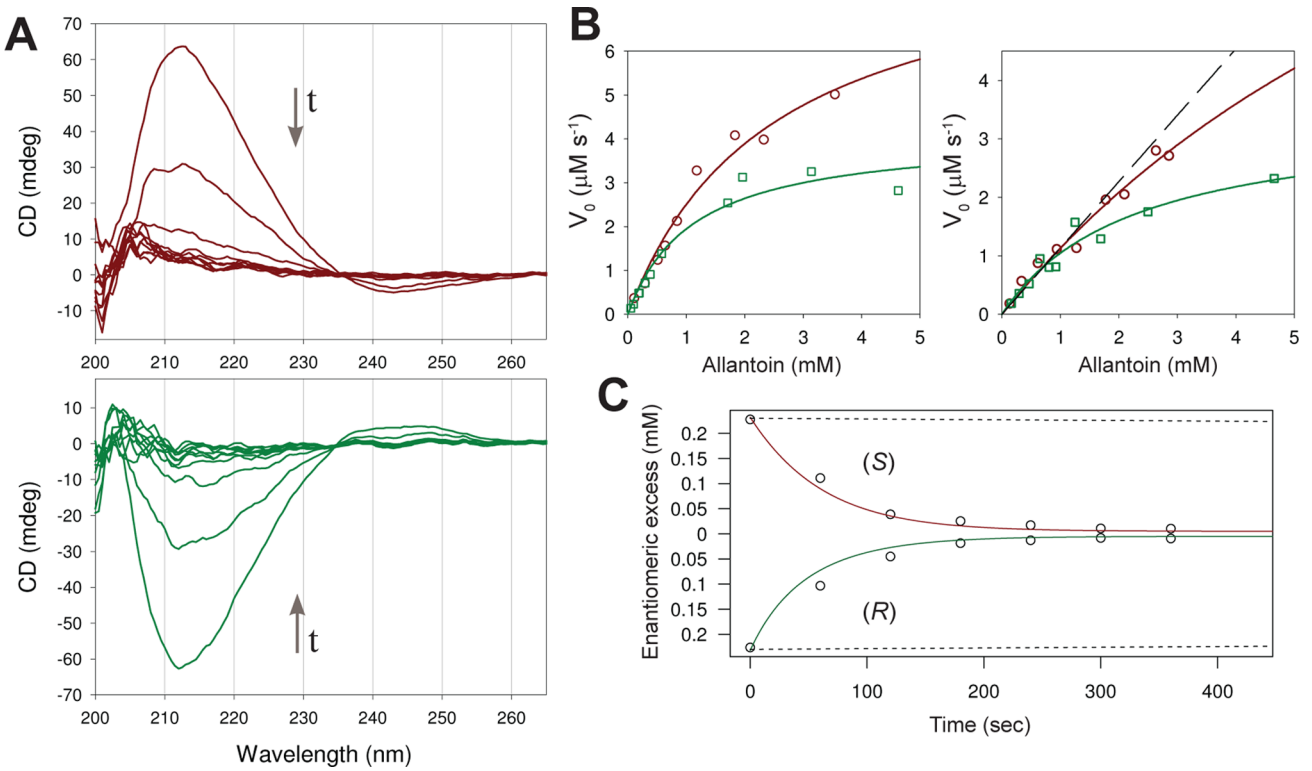



**Figure 2.** Racemase activity of *P. fluorescens* AllR. (A) Time course of *S* (red) and *R* (green) allantoin CD spectra in the presence of *Pf*AllR (9 μg). Spectra were recorded every 60 s with 0.23 mM allantoin. (B) Kinetic analysis. Initial velocities were determined with different concentrations of *S* (red circles) or *R* (green circles) allantoin in the presence of 1.8 μg of wild-type *Pf*AllR (left) or the E78D mutant (right). Data were fitted with the Michaelis−Menten equation. The solid red line in the left panel is a tentative fitting yielding a $K_M$ higher than the maximal concentration used in the assay. The initial slope of the hyperbola (−−−) was used to estimate $k_{cat}/K_M$. (C) Catalyzed and uncatalyzed reaction courses. Dashed lines represent the loss of CD signal in the absence of *Pf*AllR ($k_{uncat} = 6.2 \times 10^{-5}$ s$^{-1}$); circles are experimental data points in the presence of *Pf*AllR, and red and green lines are kinetic curves[55] calculated as a function of time, initial substrate concentration, $V_{max}$ ($3.5 \times 10^{-7}$ mM × $k_{cat}$), and $K_M$.

for different substrate concentrations by linear fitting of data points collected over the first 15 s of the reaction.

**$^{13}$C and $^{1}$H NMR Assays.** Labeled [2-$^{13}$C]-(*S*)-allantoin, used for $^{13}$C NMR assays, was obtained enzymatically from [8-$^{13}$C]urate, which was synthesized by condensing 5,6-diaminouracil with [$^{13}$C]urea (Sigma) according to a described protocol.[8] Labeled urate was converted to [2-$^{13}$C]-(*S*)-allanotin in the presence of 20 units of urate oxidase from *Candida utiliz*, 28 units of catalase from *Corynebacterium glutamicum* (Fluka), 12 μg of zebrafish HIU hydrolase, and 156 μg of zebrafish OHCU decarboxylase. The conversion was performed in 1.1 mL of 0.1 M potassium phosphate [20% $D_2O$ (pH 7.4)] and monitored spectrophotometrically. After the completion of the reaction, the solution was ultrafiltered to remove proteins and then transferred into a 5 mm NMR tube. The time course of isotopic exchange of [2-$^{13}$C]-(*S*)-allantoin was monitored in the absence and presence of 9 μg of recombinant AllR. The $^{13}$C NMR spectra were proton-decoupled and recorded at 25 °C with a VARIAN Inova 600 instrument. $^{1}$H NMR assays were performed in 0.55 mL of 0.1 M potassium phosphate [98% $D_2O$ (pD 7.6)], containing 20 mM (*R,S*)-allantoin, in the absence and presence of 5.2 μg of recombinant AllR. The time courses of proton−deuterium exchange were collected at 25 °C with a Bruker AVANCE 300 instrument.

## RESULTS

**Bioinformatics Evidence of the Identification of *Pseudomonas* Allantoin Racemase.** Allantoin racemase activity was first described in *Pseudomonas* species,[5] but the genes and proteins responsible for the activity in these organisms have not been identified at the molecular level. By homology searches with the sequence of a confirmed AllR protein of *K. pneumoniae*,[18] we found several proteins that are significantly similar in *Pseudomonas* spp. The best hit in *P. fluorescens* was a protein sharing 40% sequence identity with *Kp*AllR (Figure 1A). Comparative analysis of the gene context provided additional evidence of the functional identification (Figure 1B). In *Pseudomonas*, the gene is encoded in the same operon with a gene similar to the *DAL4* allantoin transporter of *S. cerevisiae*;[34] in other organisms, the gene is found near genes similar to characterized allantoinase and allantoicase.

In the original study, the AllR activity was meticulously assayed in 17 *Pseudomonas* isolates and found to be scattered across species.[5] Several complete genomes are available of the species and strains tested for AllR activity (Table S1), allowing us to search for genotype/phenotype correlations. A whole genome analysis of orthologous genes[20] identified the candidate *Pf*AllR among the four genes (GenBank accession numbers ADE88156, PFLU0825, PFLU2141, and PFLU2444) present in AllR+ species and absent in AllR− species (Figure 1C).

Evidence from gene context and genotype/phenotype correlation would have allowed the functional identification of the gene even in the absence of a characterized homologue. Paradoxically, the very examination of the homology evidence calls this identification into question. The *Pseudomonas* protein lacks a cysteine residue (C79 in the *Kp*AllR sequence) that is essential for the two-base catalytic mechanism of *Kp*AllR. The


DOI: 10.1021/acs.biochem.6b00881
*Biochemistry* XXXX, XXX, XXX−XXX

**Table 1. Kinetic Constants for Wild-Type and Mutant *Pf*AllR**

| | *Pf*AllR | | Glu78Asp | |
|---|---|---|---|---|
| | (*S*)-allantoin | (*R*)-allantoin | (*S*)-allantoin | (*R*)-allantoin |
| $k_{cat}$ ($s^{-1}$)[a] | 129 ± 14 | 57 ± 6 | not determined | 43 ± 7 |
| $K_M$ ($mM^{-1}$)[a] | 2.5 ± 0.6 | 1.1 ± 0.3 | not determined | 2.3 ± 0.6 |
| $k_{cat}/K_M$ ($M^{-1}$ $s^{-1}$) | (5.2 ± 1.8) × $10^4$ | (5.2 ± 1.9) × $10^4$ | (1.6 ± 1.1) × $10^4$ | (1.9 ± 0.8) × $10^4$ |

[a]Kinetic parameters were determined by nonlinear fitting of the data reported in Figure 2.

residue is substituted with glycine in the candidate *Pf*AllR sequence (see Figure 1A), while the more conservative Cys79Ser substitution was found to completely abolish the enzyme activity.[18] The Cys79Gly substitution is observed in an ungapped block of the alignment in multiple homologous sequences, thus ruling out the possibility of sequencing or misalignment errors.

**Recombinant Expression and Activity of *Pf*AllR.** In the presence of contrasting evidence of bioinformatics analysis, we decided to test the activity of the isolated protein. The candidate gene was cloned from a *P. fluorescens*-type strain in *E. coli*, and the corresponding untagged protein was overproduced and purified to apparent homogeneity by ion exchange and gel filtration chromatography. The size of the protein as determined by gel filtration (~55 kDa) is consistent with a dimeric quaternary structure in solution. The purified protein, assayed through CD spectroscopy for activity on (*R*)- and (*S*)-allantoin substrates, was found to be able to catalyze the interconversion of the two enantiomers (Figure 2A), with a $k_{cat}$ of ≈130 $s^{-1}$ and a $K_M$ of 2.5 mM for $S \rightarrow R$ conversion and a $k_{cat}$ of ≈60 $s^{-1}$ and a $K_M$ of 1.1 mM for $R \rightarrow S$ conversion (Figure 2B and Table 1). Differences in the kinetic parameters are compensatory, so that the catalytic efficiency is approximately the same for the two enantiomers ($k_{cat}/K_M \approx 5 \times 10^4$ $M^{-1}$ $s^{-1}$) as reflected by the similar reaction progression curves (Figure 2C). Compared to that of the uncatalyzed reaction, the rate of allantoin racemization is enhanced by 6 orders of magnitude by the enzyme (Table 2).

**Table 2. Enhancement of Macroscopic and Microscopic Reaction Rate Constants of Allantoin Racemization by *Pf*AllR**

| | loss of optical activity[a] | proton exchange[b] | C2−C7 isotopic exchange[c] |
|---|---|---|---|
| $k_{cat}$ ($s^{-1}$) | 129 | 165 | ≤7.4 × $10^{-5}$ |
| $k_{uncat}$ ($s^{-1}$) | 6.2 × $10^{-5}$ | 1.1 × $10^{-5}$ | 7.4 × $10^{-5}$ |
| rate enhancement | 2 × $10^6$ | 1 × $10^7$ | ≤1 |

[a]Loss of CD signal measured at pH 7.6 in 100 mM KP in the presence and absence of the enzyme; $k_{cat}$ is from Michaelis−Menten kinetics. [b]Proton exchange with deuterium measured at pH 7.6 in 100 mM KP and 98% $D_2O$ in the presence and absence of the enzyme; $k_{cat}$ was calculated from the initial velocity obtained with 20 mM allantoin and 5.2 μg of *Pf*AllR. [c]Isotopic exchange measured at pH 7.4 in 100 mM KP in the presence and absence of the enzyme.

**Crystal Structure of *Pf*AllR.** The crystal structure of *P. fluorescens* allantoin racemase has been determined in two different crystal forms (cubic and orthorhombic) at 2.10 and 2.15 Å resolution, respectively (Table 3). The overall folding of the monomer (Figure 3A) is similar to that of other proteins of the Asp/Glu racemase superfamily, two domains, each containing a four-stranded parallel β-sheet sandwiched between two pairs of α-helices. The highest degree of structural similarity is with the AllR enzyme from *K. pneumoniae*.[18] The root-mean-square deviations (rmsds) between equivalent Cα atoms of one monomer of our structures compared to one of the *K. pneumoniae* [Protein Data Bank (PDB) entry 3QVJ] are 1.30 and 1.29 Å for the cubic and orthorhombic forms, respectively.

The monomeric structure of *Pf*AllR in the two crystal forms is virtually the same, the rmsd between the Cα atoms of the dimer of the cubic form and one of the three dimers present in the orthorhombic form being 0.41 Å. In the orthorhombic form, the electron density is clearly visible from residue 1 (Met) to 236, except for residues from position 144 to 153 in chains A and G (chains of the three dimers are labeled A−B, C−D, and F−G). In the same two chains, some disorder is also present in the region of residues 180−190. The latter area includes Cys 180, one of the key residues in the active site. In the cubic form, the area of residues 180−190 is well-defined, including Cys 180. There is extra electron density that corresponds to the sulfur atom of this latter residue, suggesting that the cysteine sulfur is oxidized (possibly to sulfinic acid). The region of residues 144−151 is characterized by poor electron density and high *B* factors. This flexible region could have a role in controlling the access of the substrate at the active site.

The *P. fluorescences* allantoin racemase presents the same quaternary structure in the crystal state of the enzyme from *K. pneumoniae*, six monomers arranged as a trimer of dimers around a 3-fold axis (Figure 3B,C). The latter is a crystallographic axis in the case of the cubic crystal form, and a noncrystallographic symmetry axis in the case of the orthorhombic form. An analysis performed with the PISA server indicates that the enzyme has a hexameric assembly in the crystal. The area buried during formation of the complex is 24150 $Å^2$ over a total accessible surface of 41960 $Å^2$. This is in perfect agreement with the quaternary assembly observed in *K. pneumoniae* allantoin racemase

However, a striking difference between the two proteins is observed at the active site, identified by superimposing the protein backbone of *Pf*AllR onto the *Kp*AllR structure with a bound allantoin molecule (Figure 3D). With respect to the equivalent Asp 82 residue of *Kp*AllR, the Glu 78 side chain of *Pf*AllR points in the opposite direction, toward the substrate. The carboxylic group of Glu 78 nearly occupies the space taken by the thiol group of Cys 79 in *KpAllR*, a residue substituted with Gly in the *P. fluorescens* protein (see Figure 1A). The other residues that line the active site cavity, including the other catalytic cysteine, are conserved, in terms of both nature and conformation. The substitution of a carboxylic group for a thiol group at the protein active site could influence the mechanism of allantoin racemization in *Pf*AllR.

**Mechanisms of Allantoin Racemization.** The atomic structure of *Pf*AllR suggested that a glutamate residue (E78) could functionally replace the substituted cysteine in catalyzing proton exchange at the allantoin chiral center. Direct evidence of this mechanism of catalysis can be provided by NMR


DOI: 10.1021/acs.biochem.6b00881
Biochemistry XXXX, XXX, XXX−XXX

**Table 3. Data Collection and Refinement Statistics**

| | orthorhombic | cubic |
|---|---|---|
| wavelength (Å) | 0.97372 | 1.0000 |
| space group | $P2_12_12_1$ | $P2_13$ |
| cell parameters (Å) | $a$ = 60.22, $b$ = 142.32, $c$ = 146.02 | $a = b = c$ 109.72 |
| no. of monomers in the asymmetric unit | 6 | 2 |
| resolution (Å) | 39.78−2.10 (2.21−2.10) | 77.58−2.15 (2.27−2.15) |
| $R_{sym}$ or $R_{merge}$ | 0.068 (0.329) | 0.080 (0.212) |
| $\langle I/\sigma(I)\rangle$ | 11.4 (3.3) | 19.5 (12.1) |
| completeness (%) | 99.4 (96.9) | 97.4 (86.1) |
| redundancy | 3.8 (3.7) | 9.4 (9.2) |
| | Refinement | |
| no. of reflections | 73475 | 23619 |
| $R_{work}/R_{free}$ | 0.188/0.233 | 0.205/0.264 |
| no. of atoms | | |
| protein | 10557 | 3505 |
| water | 877 | 288 |
| rmsd | | |
| bond lengths (Å) | 0.008 | 0.016 |
| bond angles (deg) | 1.2 | 1.9 |
| Ramachandran plot (%) | | |
| favored | 94.9 | 92.2 |
| allowed | 4.4 | 7.3 |
| generously allowed | 0.6 | 0.5 |
| outliers | 0 | 0 |
| overall $G$ factor | 0.2 | 0.0 |

analysis. Allantoin can undergo racemization by two mechanisms that can be separately monitored by $^{13}C$ and $^{1}H$ NMR: formation and hydrolysis of a bicyclic symmetrical intermediate and proton exchange with the solvent (Figure 4A and Figures S1 and S2). The first mechanism, which involves the scrambling between C2 and C7 atoms (159.4 and 159.7 ppm, respectively, at pH 7.4), was monitored by following the $^{13}C$ signals of (*S*)-allantoin labeled at C2; the second mechanism was followed by monitoring the exchange of the C5 proton (5.27 ppm at pH 7.6) with solvent in deuterated water (Figure 4B). The velocity of racemization through the bicyclic intermediate was found to be independent of the presence of the enzyme (Figure 4C). According to previous results,[10] the velocity of proton exchange was slower than that of C2−C7 scrambling, but this velocity was increased by 7 orders of magnitude by the enzyme (Figure 4D and Table 2).

To gain evidence of the involvement of E78 in the proton exchange mechanism catalyzed by *Pf*AllR, this residue was replaced by site-directed mutagenesis with Asp (E78D) or Gln (E78Q). The mutant proteins were produced and purified using the same procedures that were used for the wild type and assayed in the racemization reaction. The mutation with a shorter carboxylic side chain (E78D) resulted in a diminished catalytic activity (Figure 4E) that affected $k_{cat}$ and $K_M$ for both enantiomers (see Figure 2A and Table 1). The mutation with the amide group (E78Q) resulted in the loss of catalytic activity (Figure 4E). These results are consistent with the conservation of the E78 residue in proteins belonging to the same orthologous group as *Pf*AllR: a glutamate residue is present at equivalent positions except for an aspartate in a few cases.

On the basis of the evidence presented above and the *Pf*AllR structure, we conclude that the mechanism of racemization of *Pf*AllR (Figure 4F) involves Glu 78 acting as a base for the abstraction of the C5 proton of (*S*)-allantoin and Cys 180 transferring a proton on the opposite side of the allantoin ring. In the stereoinversion of (*R*)-allantoin, proton abstraction is performed by Cys 180 and reprotonation by Glu 78. The charge of the oxyanion in the deprotonated intermediate[18,19] could by stabilized through H-bond interactions formed by the backbone N−H group of Gly 181 and Tyr 76 that in the *Pf*AllR structure are in the proximity of the carbonyl oxygen on C5 of the docked substrate (see Figure 3D).

**Origin and Conservation of the Catalytic Mechanism of Allantoin Racemase.** Experimental evidence indicates that AllR enzymes have been selected during evolution for the catalysis of allantoin racemization through a mechanism that is kinetically less favorable than an alternative one; *Pf*AllR has conserved this mechanism in spite of the substitution of a catalytic residue. To gain insights into the evolutionary origin of the catalytic mechanism of AllR, we examined its phylogenetic relationships and distribution within the Asp/Glu racemase superfamily (Figure 5). For this analysis, we identified, through hidden Markov model (HMM) searches, 2702 Asp/Glu racemase members in a set of 1700 reference complete genomes.

The protein phylogenetic tree shows four main groups; a tentative function can be assigned to groups or subgroups on the basis of the presence of characterized proteins. The most numerous group (GLU) is represented by proteins assigned to glutamate racemase by similarity with the structural and functional characterized enzymes from *Aquifex pyrophilus* (PDB entry 1B74), *Helicobacter pylori* (PDB entry 2JFY), *E. coli* (PDB entry 2JFN), and *Bacillus anthracis* (PDB entry 2DWU).[35−37] A second group (ASP) phylogenetically close to Glu racemase is represented by aspartate racemase and includes the characterized enzyme from the archaeon *Pyrococcus horikoshii* (PDB entry 1JFL).[38] Both glutamate and aspartate racemase exhibit strict conservation of the two catalytic cysteines (Figure 5). By contrast, in numerous sequences of a distinct subgroup allied to aspartate racemase, the first cysteine


DOI: 10.1021/acs.biochem.6b00881
Biochemistry XXXX, XXX, XXX−XXX

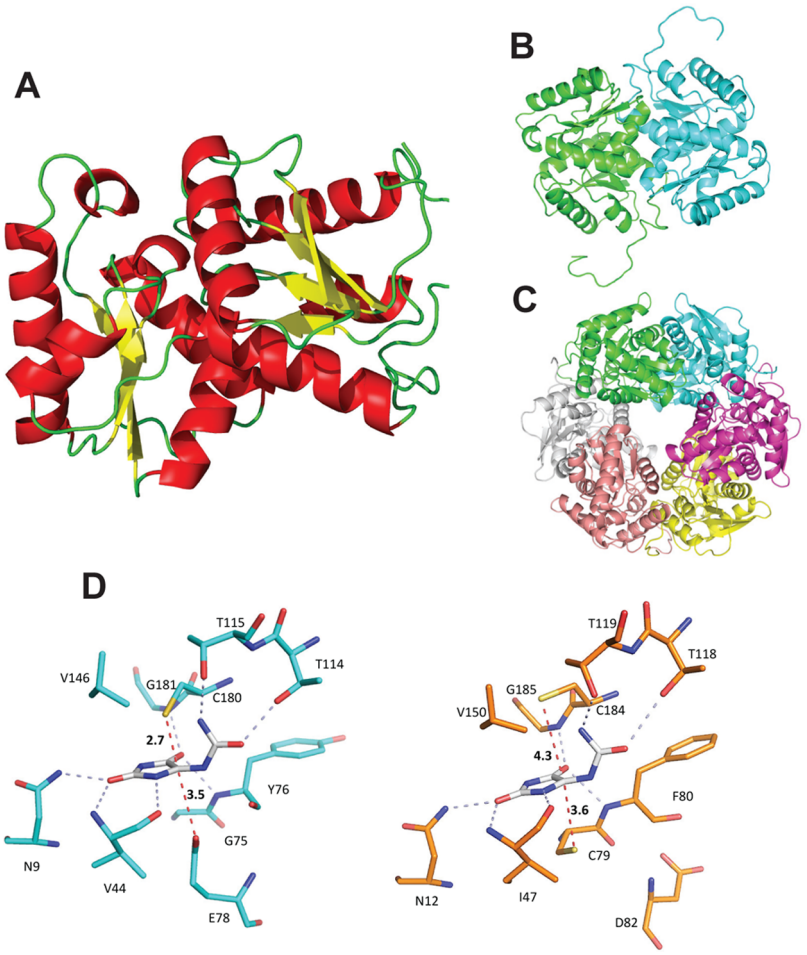



**Figure 3.** Crystal structure of *Pf*AllR. (A) Monomer in cartoon representation. (B) Dimeric organization in the cubic crystal form. The hexamer is generated by the crystallographic 3-fold axis. (C) Hexameric organization in the orthorhombic crystal form. (D) Comparison of the active site of *Pf*AllR in the cubic crystal form (left, cyan carbon atoms) and *Kp*AllR (right, orange carbon atoms, PDB entry 3QVJ). Enolic allantoin (gray carbon atoms) was docked by superimposing the backbone chain of the two protein structures onto the protein−substrate complex of the *Kp*AllR double mutant (PDB entry 3QVK). Hydrogen bonded contacts are indicated by gray dashed lines. Distances from the putative catalytic dyads to the allantoin chiral center are reported in angstroms.

is often substituted with threonine or alanine. A member of this subgroup, named EcL-DER (PDB entry 5ELL), has recently been reported to exhibit racemase activity with the L enantiomers of Glu and Asp.[39] A third group (MAL) includes characterized maleate isomerase from *Nocardia farcinica* (PDB entry 2XEC) and *Pseudomonas putida* (PDB entry 4FQ7)[40,41] and malonate decarboxylase from *Bordetella brochiseptica* (PDB entry 3DG9).[42] The first cysteine of the catalytic dyad is typically conserved in the maleate isomerase subgroup but not in the malonate decarboxylase subgroup. Finally, a fourth group (HYD) includes sequences characterized as hydantoin racemase (GenBank entries AAQ93382 and BD181026)[43,44] and allantoin racemase (*Kp*AllR and *Pf*AllR). The two enzymes catalyzing the racemization of allantoin (5-ureido hydantoin) form two distinct subgroups that are intermixed by sequences characterized as hydantoin racemase. A common characteristic of hydantoin and allantoin racemase is the presence of an acidic residue at position +3 with respect to the position occupied by the N-terminal cysteine of the dyad; in the *Pf*AllR subgroup, this residue is glutamate, while the cysteine is substituted with a conserved glycine (Figure 5).

By constructing specific HMM profiles, we observed the occurrence of the four main phylogenetic groups of the superfamily (GLU, ASP, MAL, and HYD) in 143 completely sequenced organisms classified in the Tree of Life.[45] The gene distribution (Figure S3) shows that the Asp/Glu superfamily is mostly represented by bacterial sequences. Among eukaryotes, genes belonging to the ASP group are found in plants, and genes belonging to the HYD group are found in fungi (including the putative allantoin racemase *Dcg1p* of *S. cerevisiae*), while members of the superfamily are absent in animals. Indeed, in previous work, the presence of allantoin racemase sequences in the genomes of nematodes led to the identification of a contaminant bacterial genome.[46] Genes belonging to the ASP, MAL, and HYD groups are found in some archeae. However, in the phylogenetic tree, these archaeal sequences cluster with bacterial sequences, suggesting that they could derive from lateral gene transfer.

Overall, the data support an origin of the Asp/Glu superfamily in the common ancestor of all bacteria, with the GLU group as the founding member of the superfamily. This is consistent with the presence of orthologous genes of the GLU group in the vast majority of bacteria, and with the fundamental role of glutamate racemase in providing the D-glutamate building block of peptidoglycan. This defining feature of the bacterial cell is thought to be already present in the last


DOI: 10.1021/acs.biochem.6b00881
*Biochemistry* XXXX, XXX, XXX−XXX

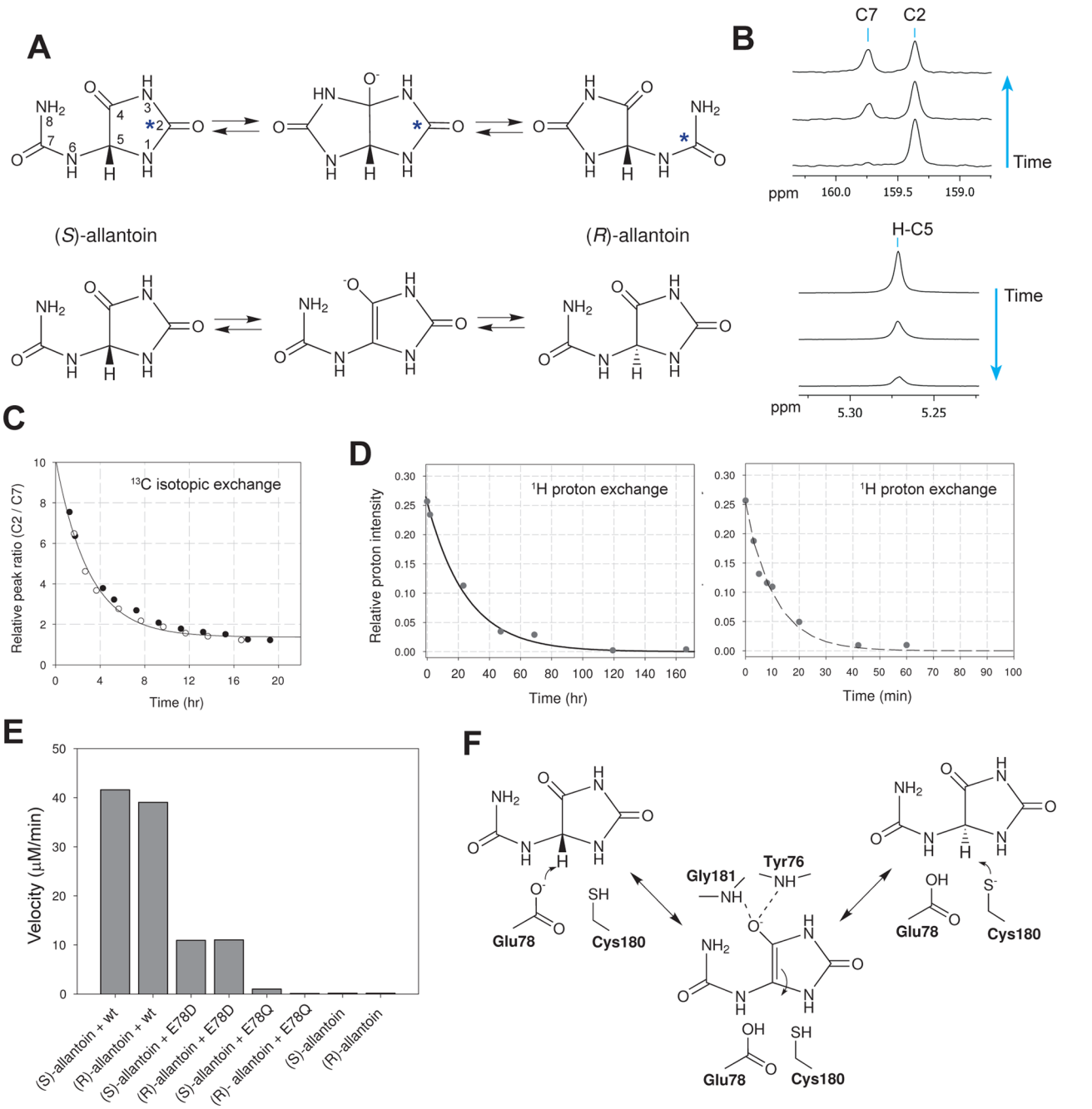



**Figure 4.** Uncatalyzed and catalyzed allantoin racemization mechanisms. (A) Alternative mechanism of allantoin racemization. The labeled carbon atom in the (*S*)-allantoin molecule is denoted with an asterisk. (B) Examples of the time-dependent isotopic exchange between C2 and C7 (top) and H−C5 proton−deuterium exchange (bottom) as monitored by $^{13}C$ and $^{1}H$ NMR. (C) Kinetics of isotopic exchange between C2 and C7 of allantoin at pH 7.4 in the absence (○) and presence (●) of *Pf*AllR. (D) Kinetics of C5 proton exchange in 98% $D_2O$ (pH 7.6) in the absence (right) and presence (left) of *Pf*AllR. (E) Comparison of the activity toward the *S* and *R* enantiomers of allantoin (0.2 mM) for the wild-type and Glu 78 mutants (1.8 μg). (F) Proposed mechanism of allantoin racemization catalyzed by *Pf*AllR.

common ancestor of bacteria.[47] A peptidoglycan cell wall has been lost by several bacterial species such as Tenericutes, *Chlamydia*, and *Wolbachia*. These species have also lost GLU racemase genes (Figure S3).

## ■ DISCUSSION

The observation of a substituted catalytic residue in protein sequence alignment is generally considered to be evidence that the protein has lost or modified the enzymatic activity. Before this conclusion can be reached, the alignment is manually inspected to rule out problems arising from misalignment of gapped regions. In the case described here, the nonconservative Cys/Gly substitution occurs in an ungapped block (see Figure 1A) and can be confidently considered as an evolutionary substitution. Consistent with data on the enzymatic activity, the atomic structure of the protein revealed a downstream residue (E78) that can functionally replace the missing catalytic residue in the two-base mechanism (see Figure 3D). These results emphasize the importance of the availability of structural information for protein functional annotation.[48]


DOI: 10.1021/acs.biochem.6b00881
*Biochemistry* XXXX, XXX, XXX−XXX

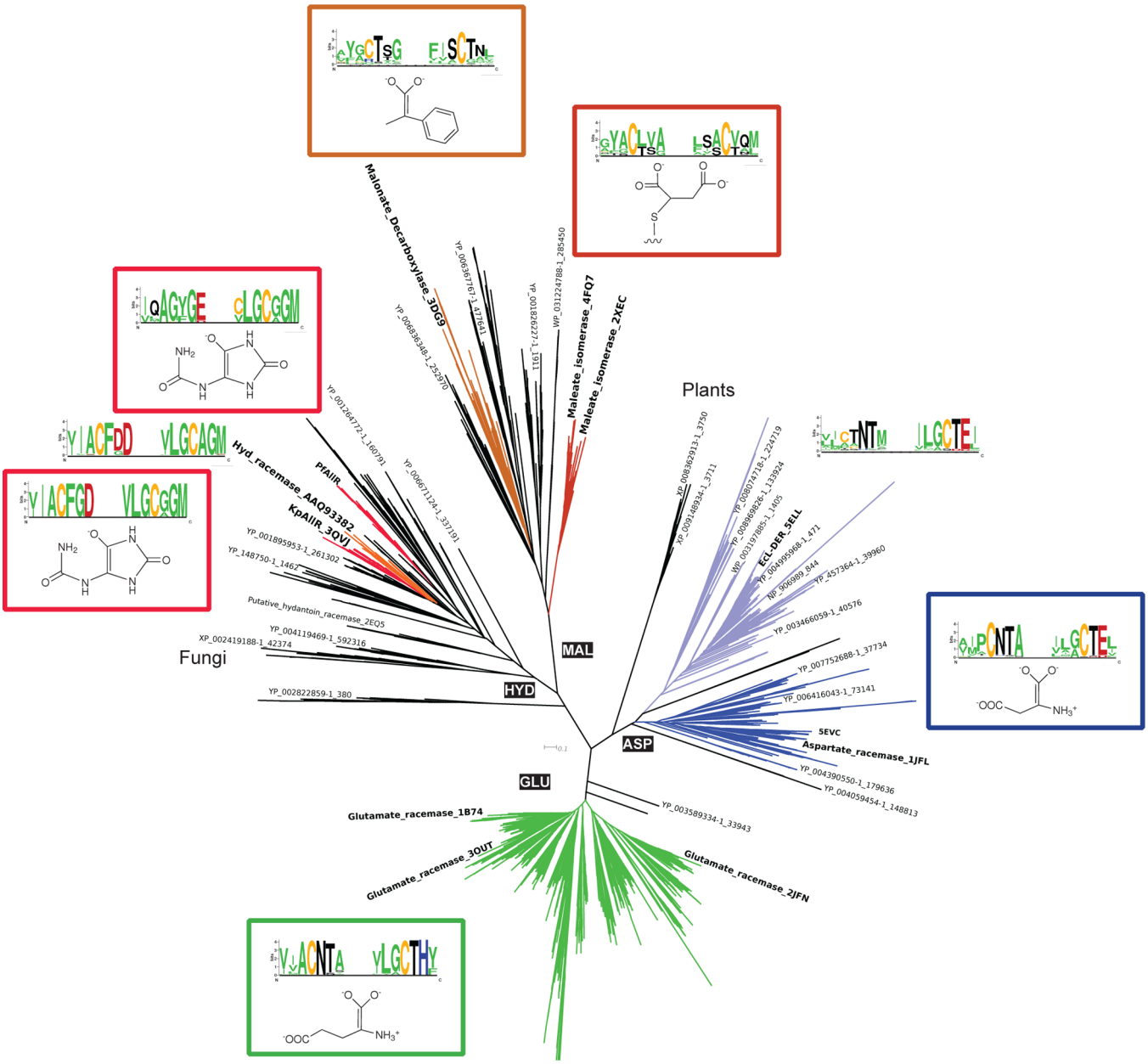



**Figure 5.** Evolution of allantoin racemase within the Asp/Glu racemase superfamily. Unrooted neighbor-joining tree of the Asp/Glu racemase superfamily. Experimentally characterized proteins are indicated in bold with PDB or GenBank accession numbers; other sequences are indicated with GenBank accession numbers followed by the taxid. Sequences from eukaryotic organisms (plants and fungi) are labeled. Tree branches assigned to functional groups based on the phylogenetic analysis are shown in different colors next to sequence logos representing residue conservation around the cysteine dyad (−3,+3). Reaction intermediates previously proposed on the basis of structural evidence are shown in colored boxes.

Changes in the identity and position of catalytic residues during divergent evolution have been observed in "mechanistically diverse superfamilies",[49] such as enolase or Nudix hydrolase. These examples, however, concern distantly related proteins catalyzing different reactions. It is uncommon to observe changes in the identity and position of catalytic residues in similar proteins catalyzing the same reaction. It is not clear what selective advantage has driven the substitution of the catalytic cysteine in allantoin racemase. Certainly, the replacement of one of the two cysteines with glutamate creates a more asymmetric active site that could allow different reactivity toward the two enantiomers. The kinetic parameters for the two enantiomers appear to be more diverse for the enzyme with the carboxyl thiol dyad (*Pf*AllR) with respect to the enzyme with the thiol dyad (*Kp*AllR), but the catalytic efficiency for the two enantiomers is similar in *Pf*AllR ($5 \times 10^4$ $M^{-1}$ $s^{-1}$) and inferior to that reported for *Kp*AllR ($6 \times 10^5$ $M^{-1}$ $s^{-1}$).[4]

The phylogeny and distribution of the Asp/Glu racemase superfamily in living organisms suggest that the original function of these proteins could be the synthesis of the D-glutamate component of the bacterial cell wall. At variance with an early origin of glutamate racemase, the enzymatic racemization of allantoin probably appeared late in evolution. For a long evolutionary time, the bioavailability of this intermediate of purine catabolism must have been limited, as in microorganisms purines are either recycled or completely degraded for carbon and nitrogen recovery.[11] Allantoin,



DOI: 10.1021/acs.biochem.6b00881
*Biochemistry* XXXX, XXX, XXX−XXX

however, gained a special role in plants and animals as the end product of purine catabolism. Some plants (e.g., tropical legumes) use allantoin for nitrogen storage and transport, and some animals (e.g., nonhominoid mammals) use allantoin for the elimination of purine nitrogen. The divergent evolution of allantoin racemase from an Asp/Glu racemase progenitor could have allowed certain organisms (particularly bacteria and fungi) to adapt to the increased availability of this compound in the environment.

The enzymatic racemization of allantoin illustrates a clear example of evolutionary choice among alternative mechanisms of catalysis (see Figure 4A). Interestingly, the mechanism selected for catalysis (proton abstraction) is kinetically less favorable than the formation of the symmetric intermediate, a reaction not exploited by the enzyme. It should be noted that in the initial phase of evolution, the presence of a mechanism of spontaneous racemization that is 10 times faster than that catalyzed by the enzyme weakened the kinetic advantage of the proto-enzyme. A proto-AllR enzyme would not have provided a selective advantage unless it had been able to increase the proton exchange rate by 1 order of magnitude.

According to the evolutionary history of Asp/Glu racemase, proton abstraction at the chiral center and stabilization of a negatively charged enedienolate intermediate[35,50] could be considered the archetypal catalytic mechanism of the superfamily (see Figure 5). This mechanism is found in glutamate racemase, aspartate racemase, and malonate decarboxylase. Variation of the common theme is found in maleate isomerase, in which the proposed intermediate is a succinyl−Cys covalent enzyme−substrate complex,[40,41] and in allantoin racemase, in which the proposed intermediate is an enolate. It should be remarked that the stabilization of the negative charge of the bicyclic intermediate could in principle allow an enzyme to catalyze racemization through the alternative reaction mechanism (see Figure 4A). However, QM/MM calculations of the *Kp*AllR reaction suggest that the allantoin substrate binds the enzyme with an extended *trans*-ureido group,[19] a conformation that is unfavorable for the nucleophilic attack of ureic nitrogen on the C4 carbonyl group and the formation of the bicyclic intermediate.

The origin and evolution of enzyme catalysis are problems of significant interest.[51−53] When a catalytic mechanism of an enzymatic reaction is studied, a relevant question is why the enzyme developed that particular catalytic strategy. In many cases, the answer is that a protein using that catalytic strategy already existed that could be adapted for a novel substrate. As shown by the case presented here, this answer implies that alternative, viable mechanisms of catalysis that have not been selected by evolution do exist. Thus, while conservation of structural and mechanistic aspects of catalysis appears to dominate the evolution of enzyme families, the choice of a particular catalytic mechanism for an enzyme can easily depend on its ancestry. The fact that novel functions are created by modification of old functions rather than by *de novo* invention is a well-established notion in biology as epitomized by Jacob's statement "Nature is a tinkerer, and not an inventor".[54] Although it is generally thought that this notion also applies to the evolution of catalytic mechanisms, it is not easy to find examples in which the choice between alternative reaction mechanisms can be readily explained by the ancestry of the protein responsible for the catalysis. Allantoin racemase provides circumstantial evidence that the evolutionary origin of an enzyme can be the decisive factor in the choice of a particular catalytic mechanism.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.6b00881.

Genome accessions of the *Pseudomonas* species previously tested for AllR activity (Table S1), primers for *Pf*AllR amplification and mutagenesis (), $^{13}$C NMR data for the enzyme-catalyzed and uncatalyzed C2−C7 isotopic exchange of allantoin (Figure S1), $^{1}$H NMR data for the enzyme-catalyzed and uncatalyzed proton−deuterium exchange of allantoin (Figure S2), and organism distribution of the different groups of the Asp/Glu racemase superfamily (Figure S3) (PDF)

### Accession Codes

The sequence of *P. fluorescens* AllR has been submitted to GenBank as accession number ADE88156. Atomic coordinates and structure factors have been deposited in the Protein Data Bank for immediate release as entries 5FLD and 5LG5 for the cubic and orthorhombic crystal forms, respectively.

## ■ AUTHOR INFORMATION


### Corresponding Author

*Department of Life Sciences, University of Parma, Parco Area delle Scienze 23/a, 43214 Parma, Italy. E-mail: riccardo.percudani@unipr.it. Telephone: +39 0521 905140. Fax: +39 0521 905151.

### Present Addresses

[‖]L.C.: Department of Biology, University of Padova, Padova, Italy.

[⊥]I.R.: Department of Biomedicine, Biotechnology and Translational Research, University of Parma, Parma, Italy.

### Author Contributions

L.C. and I.R. contributed equally to this work.



### Funding

This work was supported by Telethon Grant GGP13149 to R.P.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Claudia Folli for discussion and help with the experiments.


## ■ REFERENCES


(1) Ramazzina, I., Folli, C., Secchi, A., Berni, R., and Percudani, R. (2006) Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nat. Chem. Biol.* 2, 144−148.

(2) Cendron, L., Berni, R., Folli, C., Ramazzina, I., Percudani, R., and Zanotti, G. (2007) The structure of 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase provides insights into the mechanism of uric acid degradation. *J. Biol. Chem.* 282, 18182−18189.

(3) Kim, K., Park, J., and Rhee, S. (2007) Structural and functional basis for (S)-allantoin formation in the ureide pathway. *J. Biol. Chem.* 282, 23457−23464.

(4) French, J. B., and Ealick, S. E. (2010) Structural and mechanistic studies *on Klebsiella pneumoniae* 2-Oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline Decarboxylase. *J. Biol. Chem.* 285, 35446−35454.




DOI: 10.1021/acs.biochem.6b00881
*Biochemistry* XXXX, XXX, XXX−XXX

(5) Van Der Drift, L., Vogels, G. D., and Van Der Drift, C. (1975) Allantoin racemase: A new enzyme from *Pseudomonas* species. *BBA - Enzymol. 391*, 240−248.
(6) Ramazzina, I., Cendron, L., Folli, C., Berni, R., Monteverdi, D., Zanotti, G., and Percudani, R. (2008) Logical identification of an allantoinase analog (puuE) recruited from polysaccharide deacetylases. *J. Biol. Chem. 283*, 23295−23304.
(7) Kim, K., Kim, M., Chung, J., Ahn, J. H., and Rhee, S. (2009) Crystal Structure of Metal-Dependent Allantoinase from *Escherichia coli*. *J. Mol. Biol. 387*, 1067−1074.
(8) Kahn, K., Serfozo, P., and Tipton, P. A. (1997) Identification of the true product of the urate oxidase reaction. *J. Am. Chem. Soc. 119*, 5435−5442.
(9) Patrício, E. S., Prado, F. M., da Silva, R. P., Carvalho, L. A. C., Prates, M. V. C., Dadamos, T., Bertotti, M., Di Mascio, P., Kettle, A. J., and Meotti, F. C. (2015) Chemical Characterization of Urate Hydroperoxide, A Pro-oxidant Intermediate Generated by Urate Oxidation in Inflammatory and Photoinduced Processes. *Chem. Res. Toxicol. 28*, 1556−1566.
(10) Kahn, K., and Tipton, P. A. (2000) Kinetics and mechanism of allantoin racemization. *Bioorg. Chem. 28*, 62−72.
(11) Vogels, G. D., and Van der Drift, C. (1976) Degradation of purines and pyrimidines by microorganisms. *Bacteriol. Rev. 40*, 403−468.
(12) Muratsubaki, H., Satake, K., and Enomoto, K. (2006) Enzymatic assay of allantoin in serum using allantoinase and allantoate amidohydrolase. *Anal. Biochem. 359*, 161−166.
(13) Cairo, M. S., and Bishop, M. (2004) Tumour lysis syndrome: New therapeutic strategies and classification. *Br. J. Haematol. 127*, 3−11.
(14) Roche, A., Pérez-Dueñas, B., Camacho, J. A., Torres, R. J., Puig, J. G., García-Cazorla, A., and Artuch, R. (2009) Efficacy of Rasburicase in Hyperuricemia Secondary to Lesch-Nyhan Syndrome. *Am. J. Kidney Dis. 53*, 677−680.
(15) 's-Gravenmade, E. J., Vogels, G. D., and van Pelt, C. (1969) Preparation, properties and absolute configuration of (−)-allantoin. *Recl. des Trav. Chim. des Pays-Bas 88*, 929−939.
(16) Wong, S., and Wolfe, K. H. (2005) Birth of a metabolic gene cluster in yeast by adaptive gene relocation. *Nat. Genet. 37*, 777−782.
(17) Pope, S. D., Chen, L. L., and Stewart, V. (2009) Purine utilization by *Klebsiella oxytoca* M5al: Genes for ring-oxidizing and -opening enzymes. *J. Bacteriol. 191*, 1006−1017.
(18) French, J. B., Neau, D. B., and Ealick, S. E. (2011) Characterization of the structure and function of *Klebsiella pneumoniae* allantoin racemase. *J. Mol. Biol. 410*, 447−60.
(19) Bovigny, C., Degiacomi, M. T., Lemmin, T., Dal Peraro, M., and Stenta, M. (2014) Reaction mechanism and catalytic fingerprint of allantoin racemase. *J. Phys. Chem. B 118*, 7457−7466.
(20) Whiteside, M. D., Winsor, G. L., Laird, M. R., and Brinkman, F. S. L. (2013) OrthologueDB: A bacterial and archaeal orthology resource for improved comparative genomic analysis. *Nucleic Acids Res. 41*, D366−D376.
(21) Dehal, P. S., Joachimiak, M. P., Price, M. N., Bates, J. T., Baumohl, J. K., Chivian, D., Friedland, G. D., Huang, K. H., Keller, K., Novichkov, P. S., Dubchak, I. L., Alm, E. J., and Arkin, A. P. (2009) MicrobesOnline: An integrated portal for comparative and functional genomics. *Nucleic Acids Res. 38*, D396−D400.
(22) Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res. 22*, 4673−4680.
(23) Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: A sequence logo generator. *Genome Res. 14*, 1188−1190.
(24) Kalia, A., Rattan, A., and Chopra, P. (1999) A method for extraction of high-quality and high-quantity genomic DNA generally applicable to pathogenic bacteria. *Anal. Biochem. 275*, 1−5.
(25) Leslie, A. G. W. (2006) The integration of macromolecular diffraction data. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 62*, 48−57.
(26) Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 62*, 72−82.
(27) 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. W., 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., Sect. D: Biol. Crystallogr. 67*, 235−242.
(28) 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.
(29) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 486−501.
(30) Murshudov, G. N., Vagin, a a, Lebedev, a, Wilson, K. S., and Dodson, E. J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 247−255.
(31) Adams, P. D., Afonine, P. V., Bunkóczi, 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., Sect. D: Biol. Crystallogr. 66*, 213−221.
(32) Laskowski, R. a., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283−291.
(33) Zanotti, G., Cendron, L., Ramazzina, I., Folli, C., Percudani, R., and Berni, R. (2006) Structure of Zebra fish HIUase: Insights into Evolution of an Enzyme to a Hormone Transporter. *J. Mol. Biol. 363*, 1−9.
(34) Dorrington, R. A., and Cooper, T. G. (1993) The DAL82 protein of *Saccharomyces cerevisiae* binds to the DAL upstream induction sequence (UIS). *Nucleic Acids Res. 21*, 3777−84.
(35) Hwang, K. Y., Cho, C. S., Kim, S. S., Sung, H. C., Yu, Y. G., and Cho, Y. (1999) Structure and mechanism of glutamate racemase from *Aquifex pyrophilus*. *Nat. Struct. Biol. 6*, 422−426.
(36) Lundqvist, T., Fisher, S. L., Kern, G., Folmer, R. H. a, Xue, Y., Newton, D. T., Keating, T. a, Alm, R. a, and de Jonge, B. L. M. (2007) Exploitation of structural and regulatory diversity in glutamate racemases. *Nature 447*, 817−822.
(37) May, M., Mehboob, S., Mulhearn, D. C., Wang, Z., Yu, H., Thatcher, G. R. J., Santarsiero, B. D., Johnson, M. E., and Mesecar, A. D. (2007) Structural and Functional Analysis of Two Glutamate Racemase Isozymes from *Bacillus anthracis* and Implications for Inhibitor Design. *J. Mol. Biol. 371*, 1219−1237.
(38) Liu, L., Iwata, K., Kita, A., Kawarabayasi, Y., Yohda, M., and Miki, K. (2002) Crystal structure of aspartate racemase from *Pyrococcus horikoshii* OT3 and its implications for molecular mechanism of PLP-independent racemization. *J. Mol. Biol. 319*, 479−489.
(39) Ahn, J.-W., Chang, J. H., and Kim, K.-J. (2015) Structural basis for an atypical active site of an l-aspartate/glutamate-specific racemase from *Escherichia coli*. *FEBS Lett. 589*, 3842−3847.
(40) Chen, D., Tang, H., Lv, Y., Zhang, Z., Shen, K., Lin, K., Zhao, Y. L., Wu, G., and Xu, P. (2013) Structural and computational studies of the maleate isomerase from *Pseudomonas putida* S16 reveal a breathing motion wrapping the substrate inside. *Mol. Microbiol. 87*, 1237−1244.
(41) Fisch, F., Fleites, C. M., Delenne, M., Baudendistel, N., Hauer, B., Turkenburg, J. P., Hart, S., Bruce, N. C., and Grogan, G. (2010) A covalent succinylcysteine-like intermediate in the enzyme-catalyzed transformation of maleate to fumarate by maleate isomerase. *J. Am. Chem. Soc. 132*, 11455−11457.
(42) Okrasa, K., Levy, C., Hauer, B., Baudendistel, N., Leys, D., and Micklefield, J. (2008) Structure and mechanism of an unusual malonate decarboxylase and related racemases. *Chem. - Eur. J. 14*, 6609−6613.


DOI: 10.1021/acs.biochem.6b00881
*Biochemistry* XXXX, XXX, XXX−XXX

(43) Martínez-Rodríguez, S., Las Heras-Vázquez, F. J., Mingorance-Cazorla, L., Clemente-Jiménez, J. M., and Rodríguez-Vico, F. (2004) Molecular Cloning, Purification, and Biochemical Characterization of Hydantoin Racemase from the Legume Symbiont *Sinorhizobium meliloti* CECT 4114. *Appl. Environ. Microbiol. 70*, 625−630.
(44) Suzuki, S., Onishi, N., and Yokozeki, K. (2005) Purification and characterization of hydantoin racemase from *Microbacterium liquefaciens* AJ 3912. *Biosci., Biotechnol., Biochem. 69*, 530−536.
(45) Ciccarelli, F. D., Doerks, T., von Mering, C., Creevey, C. J., Snel, B., and Bork, P. (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science 311*, 1283−1287.
(46) Percudani, R. (2013) A microbial metagenome (*Leucobacter* sp.) in *Caenorhabditis* whole genome sequences. *Bioinform. Biol. Insights 7*, 55−72.
(47) Errington, J. (2013) L-form bacteria, cell walls and the origins of life. *Open Biol. 3*, 120143.
(48) Whisstock, J. C., and Lesk, A. M. (1999) Prediction of protein function from protein sequence and structure. *Q. Rev. Biophys. 36*, 307−340.
(49) Gerlt, J. a, and Babbitt, P. C. (2001) Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies. *Annu. Rev. Biochem. 70*, 209−246.
(50) Puig, E., Garcia-Viloca, M., González-Lafont, À., Lluch, J. M., and Field, M. J. (2007) New insights into the reaction mechanism catalyzed by the glutamate racemase enzyme: PH titration curves and classical molecular dynamics simulations. *J. Phys. Chem. B 111*, 2385−2397.
(51) Hughes, A. L. (1994) The evolution of functionally novel proteins after gene duplication. *Proc. R. Soc. London, Ser. B 256*, 119−124.
(52) Näsvall, J., Sun, L., Roth, J. R., and Andersson, D. I. (2012) Real-time evolution of new genes by innovation, amplification, and divergence. *Science 338*, 384−387.
(53) O'Brien, P. J., and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol. 6*, R91−R106.
(54) Jacob, F. (1977) Evolution and tinkering. *Science 196*, 1161−1166.
(55) Schnell, S., and Mendoza, C. (1997) Closed Form Solution for Time-dependent Enzyme Kinetics. *J. Theor. Biol. 187*, 207−212.


DOI: 10.1021/acs.biochem.6b00881
*Biochemistry* XXXX, XXX, XXX−XXX