Title: Amino acid racemization in *Pseudomonas putida* KT2440

Running title: Amino acid racemization in *Pseudomonas putida* KT2440

Authors: Atanas D Radkov and Luke A Moe

Address of institution: Plant Science Building, 1405 Veterans Drive, Lexington, Kentucky, USA 40546-0312

Author's affiliation: Department of Plant and Soil Sciences, College of Agriculture, University of Kentucky, Lexington, Kentucky, USA 40546-0312

Corresponding author's email: luke.moe@uky.edu
Abstract

D-Amino acids are shown to play an increasingly diverse role in bacterial physiology, yet much remains to be learned about their synthesis and catabolism. Here we used the model soil- and rhizosphere-dwelling *Pseudomonas putida* KT2440 to elaborate on the genomics and enzymology of D-amino acid metabolism. *P. putida* KT2440 catabolized the D-stereoisomers of lysine, phenylalanine, arginine, alanine, and hydroxyproline as sole carbon and nitrogen sources. With the exception of phenylalanine, each of these amino acids was racemized by *P. putida* KT2440 enzymes. Three amino acid racemases were identified from a genomic screen and the enzymes were further characterized *in vitro*. The putative “biosynthetic” alanine racemase Alr showed broad substrate specificity, exhibiting measurable racemase activity with 9 of the 19 chiral amino acids. Among these amino acids activity was highest with lysine, and the $k_{cat}/K_M$ values with L- and D-lysine were three orders of magnitude greater than the $k_{cat}/K_M$ values with L- and D-alanine. Conversely, the putative “catabolic” alanine racemase DadX showed narrow substrate specificity, clearly preferring only the alanine stereoisomers as substrates. However, DadX did show six- and nine-fold higher $k_{cat}/K_M$ values than Alr with L- and D-alanine, respectively. The annotated proline racemase ProR of *P. putida* KT2440 showed negligible activity with either stereoisomer of the 19 chiral amino acids, but exhibited strong epimerization activity with hydroxyproline as a substrate. Comparative genomic analysis revealed differences among pseudomonads with
respect to alanine racemases that may point to different roles for these genes among closely related species.

Introduction

Of the 20 canonical proteinogenic amino acids, 19 are chiral about their $\alpha$-carbon and therefore exist in one of two spatial arrangements, referred to as the L- and D- stereoisomers. Nature has effectively selected for L-amino acids to serve as the building blocks of ribosomally synthesized peptides and as important metabolic intermediaries in the cell. Their corresponding D-enantiomers are far less prevalent in most biological systems. Nonetheless the D-enantiomer of each of the 19 amino acids is detected in biological systems (1–6), and in certain environments D-amino acids are abundant. This includes microbe-rich environments such as topsoil (7), fermented foods (8) and the rumen (9). Where free D-amino acid content is measured in bacterial culture supernatant or ethanolic extracts of freeze-dried bacterial samples, the most abundant free D-amino acid is typically D-alanine, but high concentrations of D-aspartate, D-glutamate, D-leucine and D-methionine are also noted in some species (1, 9). The D-stereoisomer of alanine comprises up to 65% of the free alanine in some samples (9), and millimolar D-alanine concentrations have been measured in culture supernatant (1). Consequently, when D-amino acids are detected in environmental samples their presence is typically attributed to bacteria. Because certain D-amino acids are essential components in bacterial peptidoglycan (e.g. D-alanine and D-glutamate) their abundance in bacterial
culture is not surprising (10). Nonetheless, the D-amino acid distribution in
culture does not necessarily match that expected simply for synthesis of
peptidoglycan (1). Recent work has shown that complex physiological processes
such as biofilm formation, peptidoglycan remodeling, and sporulation are
influenced by the presence of certain D-amino acids (1, 6, 11). For example, the
D-stereoisomers of leucine, methionine, tryptophan, and tyrosine are shown to
disassemble mature biofilms of *B. subtilis* at concentrations as low as 10nM (6).
The corresponding L-enantiomer does not have the same effect. In this case, the
bacterium is shown to synthesize the specific D-amino acid of note, but little is
known about how this is accomplished. Further, bacterial catabolism of certain
D-amino acids is noted and may be important for colonization of D-amino acid rich
environments (12, 13).

Bacterial synthesis of D-amino acids proceeds via enzymatic racemization of the
D-amino acid distribution in culture does not necessarily match that expected simply for synthesis of
D-amino acids (1). Recent work has shown that complex physiological processes
such as biofilm formation, peptidoglycan remodeling, and sporulation are
influenced by the presence of certain D-amino acids (1, 6, 11). For example, the
D-stereoisomers of leucine, methionine, tryptophan, and tyrosine are shown to
disassemble mature biofilms of *B. subtilis* at concentrations as low as 10nM (6).
The corresponding L-enantiomer does not have the same effect. In this case, the
bacterium is shown to synthesize the specific D-amino acid of note, but little is
known about how this is accomplished. Further, bacterial catabolism of certain
D-amino acids is noted and may be important for colonization of D-amino acid rich
environments (12, 13).

Bacterial synthesis of D-amino acids proceeds via enzymatic racemization of the
corresponding L-enantiomer. Amino acid racemases catalyze the
interconversion of the L- and D-enantiomers using either a PLP-dependent or
independent mechanism. The PLP-dependent alanine racemase enzyme class
has been studied extensively, owing to its potential as a target for antimicrobials
(14). These enzymes are known to be necessary for D-alanine synthesis for
peptidoglycan, as targeted disruption results in D-alanine auxotrophy (15).
However, many bacteria encode more than one annotated alanine racemase in
their genome. Catabolism of D-amino acids can also be initiated through
enzymatic racemization to form the corresponding L-amino acid, although
racemase-independent catabolic mechanisms also exist.
Pseudomonads are noted as models in ecological genomics (16, 17), pathogenesis (18, 19), and host-microbe interactions (20, 21). They may also be considered as models in D-amino acid biology. Recent work in Pseudomonas aeruginosa PAO1 and Pseudomonas putida KT2440 has revealed unique mechanisms by which D-amino acids are catabolized (12, 13, 22). Here we build on this work by assessing the D- and L-amino acid catabolic capacity of P. putida KT2440, employing a functional screen to identify genes involved in D-amino acid metabolism, characterizing enzymes involved in amino acid racemization, and identifying key differences in the genomics of D-amino acid metabolism between related pseudomonads.

Methods

Growth studies using D- and L-amino acids as the sole source of carbon and nitrogen

Sterile PG medium (23) was prepared without a carbon or nitrogen source and was supplemented with a sterile amino acid stock to a final concentration of 25mM. Three 3ml growth tests were conducted of each amino acid enantiomer. An overnight culture of Pseudomonas putida KT2440, incubated in liquid PG containing the conventional carbon and nitrogen source (0.5% glucose and 25mM ammonium sulfate (23)), was washed with sterile water and 10μl was used to inoculate each tube. Growth controls consisted of un-inoculated culture tubes. The culture tubes were maintained at 28°C with shaking and OD₆₀₀ was...
measured after 24, 48, and 72 hours (BioMate 3 Spectrophotometer, Thermo Scientific). The respective growth control for each amino acid was used as blank.

Genomic library construction and screening

Genomic DNA from *Pseudomonas putida* KT2440 was isolated using a GenElute™ bacterial genomic DNA kit (Sigma-Aldrich) according to the supplemented instructions. Genomic DNA was fragmented by hydrodynamic shearing (GeneMachines hydroshear; speed: 19 cycles: 20). It was subsequently end-repaired (DNATerminator® end-repair kit, Lucigen) and 8kb-15kb fragments were gel extracted after performing agarose gel electrophoresis (Thermo Scientific GeneJET gel extraction kit). The purified DNA was ligated into a blunt-ended pUC19 vector using T4 DNA ligase (Thermo Scientific; 2 units) and the entire volume of all reactions was transformed into electrocompetent Epi300 *E. coli* cells (Epicentre). The library was prepared by incubating the recovery in LB medium with carbenicillin selection (100µg ml⁻¹) overnight at 37°C with shaking, harvesting the cells by centrifugation, and preparing a stock in 20% glycerol maintained at -80°C (24). DNA was isolated from the recombinant genomic library by miniprep (GeneJET plasmid miniprep kit, Thermo Scientific) and used to construct four libraries in *E. coli* amino acid auxotrophic strains obtained from the Coli Genetic Stock Center (www.cgsc.biology.yale.edu; leucine (JW5807-2), lysine (JW2806-2), proline (JW0233-2), phenylalanine (JW2580-1)). Auxotrophs were in-frame, single-gene knockouts from the Keio collection (25).
Recombinant auxotrophic genomic libraries were screened for recovery of growth on minimal medium supplemented with D-proline (10mM), D-lysine (1mM), D-leucine (10mM), or D-phenylalanine (10μM) (Sigma-Aldrich) in lieu of the corresponding L-amino acid. Briefly, portions of the glycerol stock were grown in LB overnight at 37°C with shaking, after which the cultures were washed with sterile water and 50μl of 1/100 dilutions were plated on minimal PG medium containing the respective D-amino acid. Serial dilutions were also plated on LB supplemented with carbenicillin to assess the number of CFUs plated for each screen. Screen coverage was estimated using the total number of CFUs screened and the average insert size from the genomic library. The screens were developed at 28°C. Positive clones, defined as those recombinant genomic clones that conferred growth of the auxotroph in the presence of the tested D-amino acid but not in its absence, were restreaked for isolation on PG+D-amino acid medium and cultures were started for plasmid DNA isolation (GeneJET plasmid miniprep kit, Thermo Scientific). Positive clones were retransformed into a fresh auxotrophic background and verified as above. Additionally, open reading frames of interest from recovered clones (alanine racemase/DadX, PP5269, NP_747370; alanine racemase/Alr, PP3722, NP_745855; proline racemase, PP1258, NP_743418) were cloned into pUC19 (dadX F: 5'-ATATGGATCCTATGCGTCCCAGCCCGGCCTGATC-3' (BamHI), dadX R 5'-CCGCGAGCTCTCATTCGTCCGATGTAATCCTGTTG-3' (SacI); alr F: 5'-ATATGGATCCTATGCGTCCCAGCCCGGCCTGATC-3' (BamHI), alr R: 5'-CCGCGAGCTCTCATTCGTCCGATGTAATCCTGTTG-3' (SacI); proR F: 5'-
CGCCGGATCCTATGAAACAGATTCAGTCATCGAC-3' (BamHI), proR: 5'-
TAACGAATTCTCGATGCCAGCGGAAAGGTCT-3' (EcoRI)). The constructs were transformed via electroporation into proline, leucine, or lysine auxotrophic cells to verify that the identified open reading frame itself confers the growth on the D-amino acid substrate.

DNA was sequenced on an ABI 3730 instrument (Applied Biosystems) at the University of Kentucky Advanced Genetic Technology Center using a cycle sequencing kit (BigDye® Terminator v3.1 cycle sequencing kit, Applied Biosystems) to elucidate the insert DNA from each positive clone. The genomic region harboring the DNA insert from positive clones was identified using the bioinformatics resources provided by the *Pseudomonas* Genome Database (26).

**Preparation of constructs and protein purification**

Open reading frames encoding the three putative amino acid racemases described above were cloned into pET28b using the same PCR primers. The constructs were employed for the production of N-terminal Histidine-tagged protein using *E. coli* Rosetta2 (DE3) (EMD Biosciences, EMD Millipore). For gene overexpression, an overnight culture of Rosetta2 (DE3) cells containing the appropriate construct was used to inoculate 500ml LB culture, grown at 37°C with shaking until OD_{600} was approximately 0.6-0.7. The culture was induced with 0.5mM IPTG and shaken vigorously (270 rpm) for three hours at 28°C. Induced cells were harvested by centrifugation and stored at -80°C.
For protein purification, the cells were resuspended in 2ml buffer A2X (50mM HEPES pH 7.4, 200mM NaCl, 1.95mM TCEP, 10% glycerol) and lysed via sonication (ten 20s pulses at 60V with two minute breaks on ice) and the cell lysate was subjected to immobilized metal affinity column chromatography using HIS-Select® Nickel affinity gel (Sigma-Aldrich). Following loading, the column was washed with buffer A (50mM HEPES pH 7.4, 100mM NaCl, 0.97mM TCEP, 5% glycerol) supplemented with 20mM imidazole, then eluted with buffer A containing 250 mM imidazole. Fractions containing the purified protein were identified by SDS-PAGE and were concentrated and buffer exchanged into buffer A2X using centrifugal concentration (Pierce® Concentrators, 7ml/9K MWCO, Thermo Scientific). Protein concentrations were determined via the Protein Assay Dye Reagent from Bio-Rad. Purified samples were analyzed by SDS-PAGE (Precise Tris-HEPES Gels, Thermo Scientific) to assess apparent protein purity. The concentration of the purified protein was determined as above and 50μl aliquots were snap frozen in liquid nitrogen and stored at -80°C until use.

**In vitro enzyme assays**

Amino acid racemization assays were performed in 2ml Eppendorf tubes for the duration of one minute at 37°C. The total volume for each assay was 200μl and comprised the following final concentrations: 50mM HEPES (pH 7.4), 20μM Pyridoxal-5′-phosphate (PLP) (PLP was added only in the characterization of the annotated alanine racemases but was not included for the proline racemase). Reactions were initiated by the addition of purified enzyme. For substrate
specificity assays the following concentrations were used: alanine racemase/Alr - 1μM; alanine racemase/DadX – 0.7μM; proline racemase – 1.9μM. For kinetic assays the following concentrations were used: alanine racemase/Alr – 1 μM with alanine as a substrate and 1nM with lysine; alanine racemase/DadX – 0.7 μM with alanine; proline racemase – 9.5nM with cis-D-hydroxyproline and 1.9 μM with trans-L-hydroxyproline (hereafter cis-D-hypro and trans-L-hypro). To quench the enzyme reaction, 40μl of 2M HCl was added. The molar concentration of enzyme used in each reaction was 10% or less of the initial substrate concentration, and we chose a reaction duration that allowed for the conversion of 10% or less of the initial substrate to minimize the reverse reaction.

Marfey’s reagent (1-fluoro-2,4-dinitrophenyl-5-L-alanine amide, FDAA; Sigma-Aldrich) was used for the derivatization of amino acids and reverse-phase HPLC (Dionex Ultimate 3000; Waters Nova Pak® C18, 3.9mm x 150mm column) was employed to determine the concentration of each enantiomer (27). To avoid hydrolysis of Marfey’s reagent, 40μl of 2M NaOH was added to neutralize the quenched enzyme reaction prior to derivatization, followed by 100μl Marfey’s reagent (0.5% solution in acetone) and 20μl 1M NaHCO₃. The derivatization was incubated at 40°C for one hour, after which it was allowed to cool to room temperature, and diluted tenfold in 80%:20% 0.05M Triethylamine-phosphate buffer pH 3.0 (TEAP):acetonitrile (concentrated phosphoric acid was used to adjust the pH of the triethylamine solution to prepare the TEAP buffer). The resulting solution was passed through a 0.22μm filter (nonsterile hydrophobic polytetrafluoroethylene PTFE syringe filters, Tisch Scientific) and placed into an
amber glass vial prior to HPLC analysis. A gradient method was used for HPLC separation of most amino acids: starting at 80%:20% TEAP buffer (pH 3.0): acetonitrile, ramping to 70%:30% buffer: acetonitrile over 5 minutes, followed by a ramp to 50%:50% buffer: acetonitrile over 15 minutes (see Table S1 for additional HPLC gradients). The injection volume was 10 μl and the flow rate was 0.5 ml min⁻¹. Products of the derivatization were detected at 340 nm. Authentic standards of each D- and L-amino acid were used to establish retention time, and a standard curve was generated from known amino acid concentrations in the reaction buffer in the absence of enzyme.

**Phylogenetic analysis**

To determine any relationships between the identified enzymes in this study and other racemases, phylogenetic trees were constructed using software provided by Phylogeny.fr (28). MUSCLE was used for multiple alignment (no alignment curation; Run mode: default; Maximum number of iterations: 16), PhyML for tree construction (Statistical test for branch support: aLERT - SH-like; Substitution model: default), and TreeDyn for tree visualization (default parameters). Easyfig (29) was used for the genomic loci comparison and visualization. The FASTA sequences used in the tree construction were obtained from the GenBank database of the National Center for Biotechnology Information (NCBI) (30).

**Results**

**Growth profile: amino acids as the sole carbon and nitrogen source**
Other studies have described the catabolic versatility of *P. putida* KT2440, yet a comprehensive analysis of growth on D- and L-amino acids as a sole carbon and nitrogen source has not been reported (12, 31, 32). To assess the potential for racemization of D-amino acids, we performed a series of growth studies on *P. putida* KT2440 in which D- and L-amino acids were provided as the sole carbon and nitrogen source (Fig. 1). No growth was recorded on either stereoisomer of methionine, threonine, leucine, or tryptophan. In instances where only the L-stereoisomer was catabolized, aspartate, asparagine, histidine, glutamine, glutamate, and proline were the most rapidly catabolized. Rapid growth was observed with either epimer of 4-hydroxyproline, *cis*-D- and *trans*-L-hypro, or either enantiomer of alanine. The lysine and arginine enantiomers were catabolized more slowly. Marginal growth was recorded on D- and L-phenylalanine after 24 hours. However, after 72 hours, these cultures reached growth levels similar to those of the other amino acid enantiomer pairs that were more rapidly catabolized. While previous growth experiments have established that *P. putida* KT2440 catabolizes both D- and L-lysine (13, 33), *P. putida* KT2440 has not been shown previously to catabolize as sole carbon and nitrogen sources D-phenylalanine, the epimers of hypro (*cis*-D- and *trans*-L-), or D-arginine. Nonetheless, *P. putida* KT2440 can use both enantiomers of phenylalanine, of arginine, and of hypro as a sole nitrogen source (31, 32).

**Recovered genes with a potential role in amino acid racemization**
Screens were conducted in *E. coli* auxotrophs to identify mechanisms for racemization of lysine, proline, phenylalanine, and leucine. Lysine and phenylalanine auxotrophs were chosen due to the d-amino acid growth profiles of *P. putida* KT2440 (Fig. 1). The proline auxotroph was chosen due to the presence of an annotated proline racemase in the *P. putida* KT2440 genome. The leucine auxotroph was chosen for two reasons. First, no alanine auxotroph was available from the CGSC and leucine exhibits reasonable structural similarity to alanine, and a racemase that uses alanine as a substrate may also use leucine. Second, it is possible that *P. putida* KT2440 may also racemize amino acids that it cannot use as a sole carbon and nitrogen source. Figure 1 shows the four amino acids that are not catabolized as a sole carbon and nitrogen source by *P. putida* KT2440, leucine is a representative from this group. A screen on d-arginine was not conducted because of the high activity of the recovered Alr enzyme in conversion of d-arginine to l-arginine.

Screening commenced until at least 500× coverage of the *P. putida* KT2440 genome was achieved in each screen. From the recovered positive clones, DNA sequences were compared against the *P. putida* KT2440 genome using the *Pseudomonas* genome database (26) to identify ORFs encoded within each region. Because of the magnitude of the screen, identified genomic regions were typically recovered more than once, allowing us to estimate a minimal genomic region necessary for recovering the phenotype. Multiple, independent genomic regions were recovered from the d-proline, d-lysine, and d-leucine screens, while only one region was found to rescue the phenotype in the d-phenylalanine
screen. These regions (Table S2) vary in size and in many cases include more than one gene that could be responsible for the observed phenotype. A genomic region containing the putative alanine racemase \textit{air} gene (PP3722) was recovered in both the \textit{d}-lysine and \textit{d}-leucine screens. The genomic region containing the putative alanine racemase \textit{dadX} (PP5269) was recovered only in the \textit{d}-leucine screen. The putative proline racemase (PP1258) was recovered only in the \textit{d}-proline screen, and no putative racemase gene was identified from the sole genomic region that conferred growth in the \textit{d}-phenylalanine screen.

Among genomic regions in which an annotated racemase was identified, a subclone containing the racemase gene by itself conferred the same phenotype as the insert from which it originated. We concluded that the racemase genes were responsible for the observed phenotype in these cases and did not further assess other genes from these regions. Metabolism of amino acids is known to involve enzymes such as dehydrogenases, oxidoreductases, and aminotransferases (13, 22, 34), and genes annotated as such were recovered in each of the four screens (Table S2). The current study focuses on mechanisms for amino acid racemization, further work beyond the scope of this article will be necessary to establish a role for these genes in \textit{d}-amino acid metabolism in \textit{P. putida} KT2440.

**Substrate specificity of** \textit{Pseudomonas putida} KT2440 amino acid racemases
Genes encoding putative racemase activity (proline racemase PP1258, alanine racemase/Alr PP3722, alanine racemase/DadX PP5269) were subcloned into pET28b. Each enzyme was purified to homogeneity and analyzed for its activity on both stereoisomers of each of the 19 chiral amino acids (Table 1). The protein product of the gene annotated as *alanine racemase/alr* (Alr) demonstrated the broadest substrate specificity of the recovered racemases. Despite its annotation as a putative alanine racemase, substrates that conferred the highest activity (Table 1) were lysine (normalized to 100% for both d- and l-) and arginine. The protein product of the gene annotated as *alanine racemase/dadX* (DadX) demonstrated activity with both alanine stereoisomers, showed negligible activity with d-cysteine and l-serine, and exhibited no activity with the remaining amino acids. The putative proline racemase encoded by the *proR* gene (proR) did not exhibit measureable racemase activity with any of the 19 chiral amino acid enantiomers. Nonetheless, the gene did rescue the phenotype of the *E. coli* proline auxotroph strain in the presence of d-proline, which is indicative of at least minimal racemization *in vivo*. Because previous work identifies hydroxyproline epimerase activity in cell-free extract of the related *Pseudomonas putida* KT2442 (32), we assessed the activity of the putative proline racemase with four epimers of hypro. The different epimers arise from chirality about the α-carbon and the γ-carbon in the amino acid. Isomerization about the α-carbon, typical of amino acid racemization, results in interconversion of the cis-d-hydroxyproline and trans-L-hydroxyproline epimers. The conversion of either l-hydroxyproline epimer appears not to be affected by the chirality about the γ-carbon atom, described by the cis- and
trans-notations (Table 1). However, cis-D-hypro confers a much higher activity compared to trans-D-hypro suggesting a potential influence of $\gamma$-carbon chirality in the racemization of the D-hypro epimers.

**Kinetic parameters of amino acid racemases**

The alanine racemase Alr demonstrated a $K_M$ in the $\mu$M range for D-lysine while the $K_M$ for L-lysine was significantly higher (Table 2). However, the $k_{cat}$ values suggest a more rapid racemization in the direction of D-lysine formation. Nonetheless, the $k_{cat}/K_M$ values indicate a 4-fold higher overall enzymatic efficiency in the D$\rightarrow$L direction.

Considering that both Alr and DadX catalyze racemization of alanine, we established the kinetic parameters of both enzymes with each alanine enantiomer (Table 2). For the conversion of D-alanine to L-alanine, the $K_M$ value of Alr was approximately twice that of DadX, while the $k_{cat}$ value for the DadX reaction was $\sim 4\times$ greater than Alr, contributing to a $\sim 9\times$ higher $k_{cat}/K_M$ value for the DadX reaction for the conversion of D-alanine to L-alanine. In the direction of L-ala to D-ala, Alr showed a $\sim 2\times$ lower $K_M$ value and a $\sim 12\times$ lower $k_{cat}$ value, which contributes to $\sim 6\times$ higher $k_{cat}/K_M$ values for DadX.

Kinetic values for the isomerization of cis-D-hypro and trans-L-hypro (epimers about the $\alpha$-carbon) by proR are shown in Table 2. The enzyme exhibited a lower $K_M$ value for cis-D-hypro as well as a higher maximum velocity in the D$\rightarrow$L
direction, resulting in a $26\times$ greater $k_{\text{cat}}/K_M$ value in the conversion of cis-D-hypro to trans-L-hypro than in the opposite direction.

There is significant variability in reported enzyme parameters for racemases; this can likely be attributed to differences in reaction conditions. Owing to the reversible nature of the reaction, the lack of a depleting co-substrate, and the reasonably similar catalytic efficiencies in either direction, it is necessary to limit the overall reaction time and minimize substrate turnover or the reverse reaction will confound the data. The assay conditions utilized in this work were optimized for such a reaction, using low (below two µM) enzyme concentrations with one minute reaction times, ensuring that less than 10% of the total substrate from each reaction was turned over.

**Bioinformatic analysis of *P. putida* KT2440 racemase genes**

The genome of *P. putida* KT2440 possesses 894 paralogous gene domain families, which is the highest number known relative to other sequenced bacterial genomes (35). Only the genome of *P. aeruginosa* PAO1 approaches that number with 809, while other genomes possess 50% of that or fewer (17). Consequently, we assessed whether the genetic loci characterized in this study demonstrate identity and genomic synteny to those found in the genomes of other bacteria. We performed a BLASTp analysis using the deduced protein sequence of alanine racemase/Alr (Fig. 2A). The displayed branches of the constructed tree have high support values ($\geq 0.8$), suggesting a confirmation of the tree topology. We found that the *P. putida* KT2440 Alr clusters only with similar proteins from *P.*
While this is not a complete assessment of alanine racemase genes among pseudomonads, a clear delineation, with strong branch support, can be made between alr genes from *P. putida* entries and the remaining alr and dadX genes from these pseudomonads. Figure 2B also indicates strong conservation among pseudomonad dadX genes, but significant divergence among the remaining alr genes.

Analysis of alr genomic synteny among pseudomonads revealed a strong conservation of the selected genetic loci between the two *Pseudomonas putida* strains, KT2440 and GB-1 (Fig. 2C), as well as a similarly high synteny between *P. aeruginosa* and *P. protegens* Pf-5 (formerly *P. fluorescens* Pf-5). The notable difference is the lack of the alr gene in this region in the latter two species. This difference was previously noted by Yang and Lu (36) when describing the L-arginine transaminase pathway (initiated by the aruH and arul genes from Fig. 2C) in *P. aeruginosa* PAO1. The genomic synteny of the regions surrounding dadX and the proline racemase (*proR*) was also explored (Fig. S1A). We observed high synteny for the regions bordering dadX among each of the tested
pseudomonads (same species as for the alr analysis), similar to that reported by He et al. (37). Some conservation was noted among these species with regards to the proR gene (Fig. S1B). Conservation of genes involved in hypro metabolism between P. putida and P. aeruginosa was previously noted by Watanabe, et al. (31).

Discussion

In addition to elaborating on the extent of D-amino acid utilization by P. putida KT2440, we reasoned that growth profiles may give us information on which amino acids undergo racemization. For those amino acids in which both enantiomers are catabolized, the genomic screen should recover any mechanisms by which the enantiomers are directly interconverted. However, more than one mechanism exists for recovering growth using this screen. In addition to amino acid racemization, this includes deamination of the D-amino acid, producing its corresponding achiral α-keto acid, followed by stereospecific amination to produce the L-amino acid. The latter two-enzyme process is shown to be involved in conversion of D-arginine to L-arginine via coupled catabolic and anabolic dehydrogenases encoded within the dauBAR operon in P. aeruginosa (22). A genomic library clone encoding an enzyme that catalyzes only the first step of this two-step process may recover growth due to the presence of native α-keto acid transamination activity (the second step of the above process) within the E. coli screening host. The loci recovered from these genomic screens indicate that both racemization and deamination clones were identified (Table
Based on the annotated function of the recovered genes, we chose to further study the proline racemase (*proR*, PP1258), and the two alanine racemases (*alr*, PP3722; *dadX*, PP5269).

The *proR* gene product exhibited no *in vitro* racemization of proline; rather, it appears to be a 4-hydroxyproline epimerase based on its enzymatic activity. This activity is widespread among bacteria that colonize animals and plants (38, 39), owing to the abundance of trans-L-hyp in both collagen and plant cell wall proteins (40, 41). In the catabolic pathway described by Gryder and Adams for a *P. putida* isolate (42), trans-L-hyp is first isomerized to cis-D-hyp by a (then unknown) 4-hydroxyproline epimerase. The product is then further oxidatively metabolized in several steps to produce α-ketoglutarate. We report here that the *P. putida* KT2440 *proR* gene encodes the enzyme necessary to initiate this catabolic pathway via epimerization, and recommend that the current “proline racemase” designation be changed to “4-hydroxyproline epimerase”.

Interestingly, the *in vitro* reaction appears to be much more efficient in the opposite direction — that is, in conversion of the D-hyp epimer to the L-hyp epimer (Table 2). Despite the lack of *in vitro* proline racemase activity, the enzyme must provide sufficient reactivity *in vivo* to account for the relatively small amount of L-proline necessary to recover growth of the proline auxotroph. Nonetheless, any proline racemase activity it exhibits in *P. putida* KT2440 would be minimal as it is not sufficient to allow for growth on D-proline. Figure 1 shows that the bacterium grows readily on L-proline.
The *P. putida* KT2440 DadX enzyme exhibits tight substrate specificity, suggesting a single role in interconversion of alanine stereoisomers. This is consistent with the activity of those DadX orthologs that have been characterized (43, 44). Often referred to as a “catabolic racemase” (45), DadX enzymes are proposed to serve a role in catabolism of L-alanine whereby L-alanine is converted to D-alanine before enzymatic dehydrogenation to form pyruvate (15). The *dadAX* genomic locus, which also encodes the D-amino acid dehydrogenase (*dadA*) used in this catabolic process (37), exhibits significant synteny among closely related pseudomonads (Figure S1 and (37)). The *P. putida* KT2440 DadX enzyme shows overall similar $k_{cat}/K_M$ values for both the $L \rightarrow D$ and $D \rightarrow L$ conversions (Table 2). Incidentally, the $k_{cat}/K_M$ values using L- and D-alanine as substrates are ~6 and 9-fold lower for the *P. putida* KT2440 Alr enzyme, respectively, largely stemming from lower $k_{cat}$ values with both L- and D-alanine. Nonetheless, in instances where both *alr* and *dadX* genes exist in the same organism, the Alr isozyme is considered as the “anabolic” alanine racemase and is attributed as the source for periplasmic D-alanine used in peptidoglycan synthesis (46). The *P. putida* KT2440 Alr exhibits minimal activity with alanine compared to other amino acids, however, raising questions about its role in D-alanine synthesis for peptidoglycan in *P. putida* KT2440. Indeed, the $k_{cat}/K_M$ with lysine is nearly 3 orders of magnitude greater than that with alanine in both reaction directions.

Based on the available sequences in GenBank, *alr*-encoded alanine racemases are more widespread than are *dadX*-encoded alanine racemases. However, the
The presence of both annotated alanine racemases in one genome is not unusual, as this is seen in a large number of Proteobacteria. Each of the 184 bacterial dadX alanine racemase gene entries in GenBank is from a member of the α-, β-, or γ-Proteobacteria, while the 1088 bacterial alr alanine racemase entries are more widely distributed (accessed 6/13/13). No uniform racemase gene nomenclature exists, however, and many genes are annotated as simply “alanine racemase”. A comparative analysis between the ten P. aeruginosa genomes and six P. putida genomes currently available at the Pseudomonas Genome Database (26) (accessed 6/12/13), revealed that each of these strains carries both the alr gene and the dadX gene. The dadAX locus is conserved among each of these strains, yet the alr locus differs according to species. With the exception of P. putida W619, each of the P. putida strains shares the same alr locus organization depicted for P. putida KT2440 in Figure 2C. The location of the P. putida KT2440 alr gene, between the arul and aruH genes (Figure 2C, first noted by Yang, et al. (36)) is noteworthy. The aruHI locus is shown to be necessary for the arginine transaminase pathway (ATA pathway) for catabolism of L-arginine in P. aeruginosa PAO1 (36).

While both P. putida KT2440 and P. aeruginosa PAO1 catabolize D-arginine, they appear to use distinctly different mechanisms to initiate this process. In P. aeruginosa PAO1, catabolism of D-arginine is initiated by conversion to L-arginine via the coupled dehydrogenases DauA and DauB of the dauBAR operon (22). DauA and DauB are both required for P. aeruginosa PAO1 to grow on D-arginine as a sole carbon and nitrogen source, implying that D-arginine is only
metabolized via conversion to L-arginine and that no other enzymatic arginine racemization mechanism exists in *P. aeruginosa* PAO1. Further, previous work demonstrates that D-lysine cannot complement a L-lysine auxotroph in *P. aeruginosa* PAO1 (47). Taken together, this indicates that no direct arginine or lysine racemization mechanism exists in *P. aeruginosa* PAO1. This indicates that the *alr* gene in *P. aeruginosa* PAO1 does not have the same substrate specificity as its *P. putida* KT2440 ortholog. Further, while the *dauBAR* locus is conserved among *P. aeruginosa* strains, it is not present in any of the *P. putida* strains from the *Pseudomonas* Genome Database.

Based on the above observations, it appears plausible that the Alr enzyme in *P. putida* KT2440 serves a role analogous to that of DauA and DauB in *P. aeruginosa* PAO1, in conversion of D-arginine to L-arginine for further catabolism. In this scenario, the *P. putida* KT2440 *alr* gene is responsible for conversion of D-arginine into L-arginine, which is then catabolized via the ATA pathway by the products of the *aruHI* operon in which the *alr* gene is incorporated (Figure 2C).

Previous work on catabolism of D- and L-lysine by *P. putida* KT2440 indicates that each proceeds through an independent pathway and that racemization of D-lysine to L-lysine prior to degradation is not required (13). While it is possible that the Alr enzyme is involved in degradation of lysine, its potential for synthesis of D-lysine and any additional role that D-lysine may play should be further explored. Catabolic routes in pseudomonads for certain L-amino acids, such as alanine (15) and 4-hydroxyproline (42), involve initial conversion from the L-stereoisomer
into the D-stereoisomer. This clearly prevents the amino acid from being incorporated into peptides and may minimize the changes in gene regulation that accompany increased pools of free amino acids in the cell (48, 49).

In environments where D-amino acids are abundant, their presence is typically attributed to the associated microbial communities (9, 50), yet outside of a few amino acids nothing is known regarding their synthesis. Indeed the D-stereoisomer of each of the 19 chiral amino acids has been identified in some capacity from either microbial cultures (51) or from microbe-enriched environments (1). The work here further establishes the pseudomonads as good models for comparative genomic analysis of bacterial metabolism and points to differences in D-amino acid metabolism between two of the most well studied members of this clade. However, despite an increased awareness of the roles that D-amino acids play in bacterial ecophysiology, much remains to be discovered concerning the synthetic mechanisms and reasons for deployment of these compounds.

Acknowledgments

This work was supported in part by grant 2011-67020-30195 from the USDA National Institute of Food and Agriculture.

References


**Figure legends**

**FIG. 1.** Ability of *Pseudomonas putida* KT2440 to use D- and L-amino acids as the sole source of carbon and nitrogen. Each color bar represents the net increase in growth during the designated time period. The absence of a particular color bar signifies lack of growth during the corresponding time period. Cultures were grown in PG liquid minimal medium. Average values are the result of three replicates.

**FIG. 2.** Phylogenetic analysis. (A) Tree of the 50 non-redundant protein sequences most identical to Alanine racemase/Alr. (B) Tree of the annotated DadX and Alr from the sequenced genomes in the *Pseudomonas* Genome Database. Distance bars represent expected proportion of substitutions per amino acid site. (C) Genomic synteny analysis on regions bordering the *alr* gene from *Pseudomonas putida* KT2440 (included are three additional genomic regions from *P. aeruginosa* PAO1, *P. protegens* Pf-5, and *P. putida* GB-1). Gene annotations based on the *Pseudomonas* Genome Database. The capped line represents actual DNA length. The gradient scale provides the percent identity for the genomic loci.
### TABLE 1. Percent normalized relative activity of racemases on D- and L-amino acids

<table>
<thead>
<tr>
<th>Purified enzyme</th>
<th>D-amino acids</th>
<th>Normalized percent activity (±SD)</th>
<th>L-amino acids</th>
<th>Normalized percent activity (±SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine racemase (Alr)</td>
<td>Lys</td>
<td>100.00±0.00</td>
<td>Lys</td>
<td>100.00±0.00</td>
</tr>
<tr>
<td></td>
<td>Arg</td>
<td>87.49±1.64</td>
<td>Arg</td>
<td>10.79±0.79</td>
</tr>
<tr>
<td></td>
<td>Met</td>
<td>11.56±0.20</td>
<td>Met</td>
<td>6.39±0.22</td>
</tr>
<tr>
<td></td>
<td>Gln</td>
<td>6.94±0.09</td>
<td>Gln</td>
<td>3.07±0.09</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>3.10±0.15</td>
<td>Ala</td>
<td>1.75±0.09</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>2.57±0.09</td>
<td>Ser</td>
<td>1.18±0.03</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>1.37±0.08</td>
<td>Asn</td>
<td>0.65±0.08</td>
</tr>
<tr>
<td></td>
<td>His</td>
<td>1.01±0.03</td>
<td>Leu</td>
<td>0.57±0.02</td>
</tr>
<tr>
<td></td>
<td>Asn</td>
<td>0.96±0.07</td>
<td>His</td>
<td>0.56±0.02</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>ND</td>
<td>Others</td>
<td>ND</td>
</tr>
<tr>
<td>Alanine racemase (DadX)</td>
<td>Ala</td>
<td>100.00±0.00</td>
<td>Ala</td>
<td>100.00±0.00</td>
</tr>
<tr>
<td></td>
<td>Cys</td>
<td>8.35±0.20</td>
<td>Ser</td>
<td>0.90±0.04</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>ND</td>
<td>Others</td>
<td>ND</td>
</tr>
<tr>
<td>Proline racemase (ProR)</td>
<td>cis-d-HyPro</td>
<td>100.00±0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>trans-L-HyPro</td>
<td>32.34±0.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cis-L-HyPro</td>
<td>32.23±1.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>trans-d-HyPro</td>
<td>22.73±0.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Percent activity values were normalized for difference in derivatization efficiency between each D- and L-enantiomer (52).
- Final assay concentration for each amino acid was 50mM, except Tyr (2.21mM) and Asp (25mM). Assays were performed in triplicate only for those amino acids that showed detectable activity.
- Others category comprises all remaining chiral proteinogenic amino acids.
- ND – No activity detected.
TABLE 2. Kinetic parameters of racemases determined on select D- and L-amino acids

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>$K_{M}$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}K_{M}$ (M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$D \rightarrow L$</td>
<td>$L \rightarrow D$</td>
<td>$D \rightarrow L$</td>
</tr>
<tr>
<td>Alanine racemase (Alr)</td>
<td>Lys</td>
<td>0.36±0.06</td>
<td>8.96±1.52</td>
<td>274.5±6.7</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>15.71±2.79</td>
<td>12.62±0.74</td>
<td>8.83±0.00</td>
</tr>
<tr>
<td>Alanine racemase (DadX)</td>
<td>Ala</td>
<td>7.73±0.63</td>
<td>24.57±2.64</td>
<td>37.47±2.67</td>
</tr>
<tr>
<td>Proline racemase</td>
<td>HyPro</td>
<td>5.26±1.12</td>
<td>15.04±4.19</td>
<td>69.63±8.60</td>
</tr>
</tbody>
</table>

The substrates with the highest normalized percent activity were chosen for kinetic analysis, as well as D-L-ala for Alr and trans-L-hypro for proline racemase.

$V_{max}$ and $K_{M}$ values were determined via non-linear regression according to the Michaelis-Menten equation, $k_{cat}$ values were calculated by dividing $V_{max}$ by the enzyme concentration.
FIG. 1. Ability of *Pseudomonas putida* KT2440 to use D- and L-amino acids as the sole source of carbon and nitrogen.

Each color bar represents the net increase in growth during the designated time period. The absence of a particular color bar signifies lack of growth during the corresponding time period. Cultures were grown in PG liquid minimal medium.

Average values are the result of three replicates. SD included on graphs.