

# Mechanistic and Evolutionary Insights from the Reciprocal Promiscuity of Two Pyridoxal Phosphate-dependent Enzymes<sup>\*[5]</sup>

Received for publication, May 22, 2016, and in revised form, July 11, 2016. Published, JBC Papers in Press, July 29, 2016, DOI 10.1074/jbc.M116.739557

Valerie W. C. Soo<sup>‡</sup>, Yuliana Yosaatmadja<sup>§</sup>, Christopher J. Squire<sup>§</sup>, and Wayne M. Patrick<sup>¶1</sup>

From the <sup>‡</sup>Institute of Natural and Mathematical Sciences, Massey University, Auckland 0632, the <sup>§</sup>School of Biological Sciences, University of Auckland, Auckland 1142, and the <sup>¶</sup>Department of Biochemistry, University of Otago, Dunedin 9054, New Zealand

Enzymes that utilize the cofactor pyridoxal 5'-phosphate play essential roles in amino acid metabolism in all organisms. The cofactor is used by proteins that adopt at least five different folds, which raises questions about the evolutionary processes that might explain the observed distribution of functions among folds. In this study, we show that a representative of fold type III, the *Escherichia coli* alanine racemase (ALR), is a promiscuous cystathionine  $\beta$ -lyase (CBL). Furthermore, *E. coli* CBL (fold type I) is a promiscuous alanine racemase. A single round of error-prone PCR and selection yielded variant ALR(Y274F), which catalyzes cystathionine  $\beta$ -elimination with a near-native Michaelis constant ( $K_m = 3.3$  mM) but a poor turnover number ( $k_{cat} \approx 10$  h<sup>-1</sup>). In contrast, directed evolution also yielded CBL(P113S), which catalyzes L-alanine racemization with a poor  $K_m$  (58 mM) but a high  $k_{cat}$  (22 s<sup>-1</sup>). The structures of both variants were solved in the presence and absence of the L-alanine analogue, (R)-1-aminoethylphosphonic acid. As expected, the ALR active site was enlarged by the Y274F substitution, allowing better access for cystathionine. More surprisingly, the favorable kinetic parameters of CBL(P113S) appear to result from optimizing the p*K*<sub>a</sub> of Tyr-111, which acts as the catalytic acid during L-alanine racemization. Our data emphasize the short mutational routes between the functions of pyridoxal 5'-phosphate-dependent enzymes, regardless of whether or not they share the same fold. Thus, they confound the prevailing model of enzyme evolution, which predicts that overlapping patterns of promiscuity result from sharing a common multifunctional ancestor.

Pyridoxal 5'-phosphate (PLP)<sup>2</sup>-dependent enzymes catalyze ~4% of all the reactions that have been classified by the Enzyme

Commission (EC) (1, 2). PLP-dependent enzymes are found in every organism, and it has been estimated that 1.0–1.5% of all genes in the genomes of free-living prokaryotes code for them (3).

Almost all known PLP-dependent enzymes catalyze reactions involving the C $\alpha$ , C $\beta$ , and C $\gamma$  atoms of amino acids or other metabolites with an amino group. Examples of the reactions catalyzed include racemization, elimination, transamination, decarboxylation,  $\beta$ - and  $\gamma$ -replacement, and retro aldol cleavage. Mechanistically, these reactions share the same first step, which is the formation of an external aldimine between the substrate and PLP (2, 4). The cofactor plays a catalytic role in labilizing one of the bonds at C $\alpha$  of the substrate, lowering the activation barrier for forming a carbanionic intermediate. The roles of the protein moiety are largely to ensure cleavage of the correct bond at C $\alpha$  and to guide the reaction of the intermediate toward the correct product (5).

Historically the PLP-dependent enzymes have been classified into five different fold types (6, 7), although this classification has recently been expanded to include two more folds that use PLP in a different mechanistic context (to stabilize radical intermediates) (2). Curiously, fold types I to V all use the same “phosphate-binding cup” to anchor PLP in the active site (8). We have argued that this is evidence for the ancient recruitment of a single cofactor-binding subdomain into the various extant folds (via non-homologous recombination) (9). This model implies partial homology between the fold types. However, for simplicity here we will refer to enzymes with different fold types as being non-homologous. Furthermore, the reaction specificities of PLP-dependent enzymes do not correlate strictly with their fold type. For example, aminotransferases that adopt fold types I and IV are known (2).

Together, these features raise two fundamental questions about the evolution of PLP-dependent enzymes. First, how have proteins with different folds evolved to control the reactivity of the cofactor? Second, what evolutionary processes or constraints explain the observed distribution of functions among folds?

In this study, we have addressed these questions by focusing on the two PLP-dependent enzymes alanine racemase (ALR) and cystathionine  $\beta$ -lyase (CBL). Previously, we showed that overexpression of ALR rescued an *Escherichia coli* strain in

\* This work was supported by a grant from the Marsden Fund (to W. M. P.) and a Rutherford Discovery Fellowship (to W. M. P.). The authors declare that they have no conflicts of interest with the contents of this article.

✂ Author's Choice—Final version free via Creative Commons CC-BY license.

[5] This article contains supplemental Fig. S1 and Tables S1 and S2.

The atomic coordinates and structure factors (codes 4WR3, 4XBJ, 4ITG, and 4ITX) have been deposited in the Protein Data Bank (<http://www.pdb.org/>).

<sup>1</sup> To whom correspondence should be addressed: Dept. of Biochemistry, University of Otago, P. O. Box 56, Dunedin 9054, New Zealand. Tel.: 64-3-479-7854; Fax: 64-3-479-7866; E-mail: wayne.patrick@otago.ac.nz.

<sup>2</sup> The abbreviations used are: PLP, pyridoxal 5'-phosphate; ALR, alanine racemase; ASKA, A complete set of *E. coli* K-12 ORF archive; AVG, L-aminoethoxyvinylglycine; CBL, cystathionine  $\beta$ -lyase; DTNB, 5,5'-dithiobis(2-nitrobenzoic acid); epPCR, error-prone PCR; IPTG, isopropyl  $\beta$ -D-1-thiogalactopyranoside; L-Ala-P, (R)-1-aminoethylphosphonic acid;

PDB, Protein Data Bank; CHES, 2-(cyclohexylamino)ethanesulfonic acid; POOL, Partial Order Optimum Likelihood.

## Reciprocal Promiscuity in Two PLP-dependent Enzymes

which the gene encoding CBL (*metC*) had been deleted (10). CBL is essential for *E. coli* to grow on minimal media, as it catalyzes the penultimate step in methionine biosynthesis. We were intrigued that ALR appeared to possess promiscuous CBL activity, as the two enzymes are not homologues (with the possible exception of their phosphate-binding cup subdomains). ALR (EC 5.1.1.1) and CBL (EC 4.4.1.8) are also unrelated with respect to the physiologically relevant substrates they recognize and the overall reactions they catalyze.

*E. coli* CBL is a homotetramer that catalyzes the formation of L-homocysteine from L-cystathionine (Fig. 1A). Each monomer adopts the type I PLP-binding fold and includes the following three domains: the N-terminal domain contributes to tetramerization; the C-terminal domain consists of an anti-parallel  $\beta$ -sheet with  $\alpha$ -helices on the solvent side; and the middle domain harbors PLP, which is found in a parallel seven-stranded  $\beta$ -sheet covered by adjacent  $\beta$ -helices (11). In the resting state, PLP forms an internal aldimine with active site residue Lys-210. The enzyme has been identified as a potential antibiotic target, as its inhibition impacts protein synthesis and also DNA synthesis (via the methyl donor, S-adenosylmethionine) in bacteria (12).

ALR interconverts L-Ala and D-Ala, with the latter being an essential constituent of the cell wall peptidoglycan. Indeed, the absence of D-Ala leads to a defective cell wall and cell death. The essentiality of ALR for cell wall biosynthesis renders it an attractive antibiotic target (13, 14). The enzyme is a homodimer, formed by head-to-tail association of two monomers. Each monomer adopts a type III PLP-binding fold (Fig. 1B). The N-terminal domain of each monomer is a  $(\beta\alpha)_8$  barrel; the C-terminal domain mainly comprises  $\beta$ -strands (15, 16). The active site is found in the core of the  $(\beta\alpha)_8$  barrel, where PLP forms an internal aldimine via a Schiff base linkage with Lys-34. An extensive and classic series of structural, mutagenic, and modeling analyses have established a two-base mechanism for ALR (16–22). In the *E. coli* enzyme, Tyr-255' (with the prime indicating a residue from the opposing monomer than Lys-34) acts as the catalytic base that deprotonates L-Ala. Lys-34 acts as the catalytic acid, protonating the carbanionic intermediate to yield D-Ala. The roles of Tyr-255' and Lys-34 are reversed during the racemization of D-Ala into L-Ala.

Here, we characterize the promiscuous CBL activity of ALR and show that it can be improved by directed evolution. We also demonstrate a novel example of reciprocal promiscuity in the absence of homology; CBL is a promiscuous and evolvable ALR. Thus, our study also highlights the short mutational routes that may have been sufficient to evolve families of homologous PLP-dependent enzymes with disparate reaction specificities.

### Results

**ALR Overexpression Rescued *E. coli* Lacking CBL Activity—***E. coli*  $\Delta metC$  is unable to grow on M9/glucose agar; however, in our hands it formed colonies on medium supplemented with 0.1 mM L-methionine after 1–2 days of incubation. This demonstrated that deleting CBL ablates methionine biosynthesis, as expected. *E. coli*  $\Delta metC$  that harbored a plasmid-encoded CBL (pCA24N-*metC*) survived in the absence of exogenous L-methionine (supplemental Table S1).

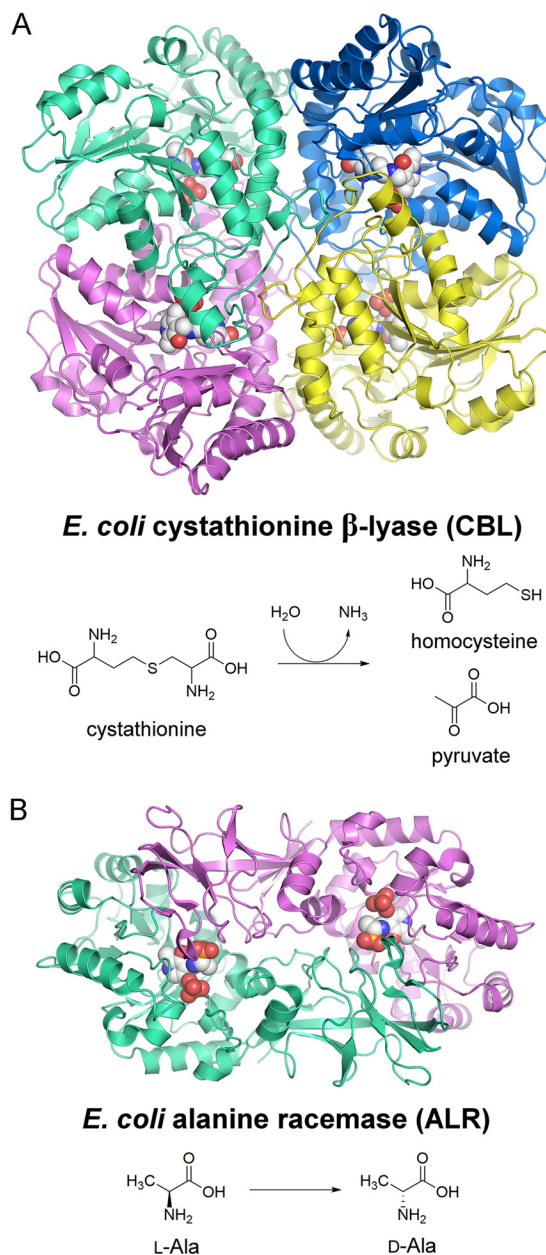


FIGURE 1. A, structure of the CBL tetramer (PDB entry 1CL1) (11) and the reaction it catalyzes in methionine biosynthesis. B, structure of the ALR dimer (PDB entry 2RJG) (15) and the reaction it catalyzes to provide D-alanine for peptidoglycan biosynthesis. In each structure, the locations of the active sites are highlighted by the PLP cofactor, which is shown in space-filling format.

Previously, we showed that ALR overexpressed from a high copy plasmid (pCA24N-*alr*) could rescue *E. coli*  $\Delta metC$ , with colony formation taking 5–6 days when expression was induced with 50  $\mu$ M isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) (10). As with all of the open reading frames (ORFs) in the ASKA collection (23), the 3' end of the *alr* insert is fused to the gene for green fluorescent protein (GFP) in pCA24N-*alr*. Therefore, we removed the GFP tag to confirm that complementation was due to ALR itself and was not a peculiarity of the ALR-GFP fusion. Cells overexpressing the GFP-less ALR did form colonies, after a shorter incubation period (3–4 days) (supplemental Table S1). Throughout this study, no growth was observed for the negative control (an *E. coli*  $\Delta metC$  clone

**TABLE 1**  
ALR mutants selected from an epPCR library for their ability to rescue *E. coli*  $\Delta metC$

Variant	Frequency	Mutation(s)
ALR-1	25	Y274F
ALR-2	1	Y274F, A335T
ALR-3	1	R266S, Y274F
ALR-4	1	M29I, Y274F
ALR-5	1	T97S, P252R, Y274F
ALR-6	1	I242V, Y274F
ALR-7	1	L241M, R263H, Y274F
ALR-8	1	V150I, Y274F

that harbored the empty vector without a gene insert, pCA24N-NoIns (24)).

This complementation assay appeared to show that ALR possessed promiscuous CBL activity. However, we also considered the possibility that ALR might rescue the methionine auxotrophy of *E. coli*  $\Delta metC$  via a metabolic bypass, rather than by catalyzing the CBL reaction. To test this, we overexpressed ALR in an *E. coli* strain lacking the enzyme that follows CBL in the biosynthetic pathway, methionine synthase. The growth of this strain (*E. coli*  $\Delta metE$ ) was tested on M9/glucose medium. If ALR overexpression had led to the growth of this strain, then this would have suggested that the cell was (somehow) rerouting metabolic flux to synthesize methionine by a previously undescribed pathway. Examples of this phenomenon have been characterized for a different metabolic pathway (25). However, we observed no growth of *E. coli*  $\Delta metE$  expressing ALR, in up to 1 week of incubation (supplemental Table S1). This demonstrated that ALR overexpression did not lead to methionine production via an alternative pathway.

Next, we attempted to measure the CBL activity of purified His<sub>6</sub>-tagged ALR by using 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) to detect the formation of homocysteine. However, no CBL activity was detected *in vitro*, at an enzyme concentration of 7  $\mu\text{M}$ , concentrations of the substrate (L-cystathionine) up to 10 mM, and at temperatures of 25, 30, or 37 °C. We could not use higher L-cystathionine concentrations due to its low solubility ( $\sim 5.6 \text{ mg}\cdot\text{ml}^{-1}$  in 0.04 N HCl). In our control reactions that lacked enzyme, we observed the spontaneous decomposition of DTNB at an average background rate of  $\leq 0.35 \mu\text{M}/\text{min}$ . This places an approximate upper limit on the CBL activity of ALR. In these assays we would have observed turnover of L-cystathionine if it had been greater than  $\sim 0.05$  per min per active site.

**Directed Evolution to Improve the CBL Activity of ALR**—Promiscuous activities are often orders of magnitude lower than native activities, and therefore, they can be exceedingly difficult to detect *in vitro* (25, 26). A simple strategy to circumvent this problem is to improve the promiscuous activity by directed evolution. The underlying assumption is that directed evolution approaches are unlikely to yield an activity that was not there before, although they are highly effective at improving weak but pre-existing functions (27).

We subjected the *alr* gene to a low mutation rate error-prone PCR (epPCR) and cloned the products into pCA24N. The resulting plasmids were used to transform *E. coli*  $\Delta metC$ , yielding a library that contained  $3.4 \times 10^7$  *alr* variants. Sequencing the *alr* gene inserts (each 1,077 bp) from 20 colonies allowed us

to use PEDEL-AA (28) to estimate the mean mutation rate at 0.8 base substitutions per allele, which corresponded to an average of 0.6 amino acid substitutions per ALR variant. The library was regrown and plated on M9/glucose with a reduced IPTG concentration (5  $\mu\text{M}$ , instead of 50  $\mu\text{M}$ ) to increase the selection stringency. Under these conditions, the parental strain (*E. coli*  $\Delta metC$  harboring pCA24N-*alr*) took 12–13 days to form colonies. Clones with higher CBL activity were selected based on their ability to form colonies more rapidly than the parental strain.

After only 2 days of incubation, we picked 32 colonies for further analysis. The growth phenotypes conferred by the selected *alr* variants were verified by isolating plasmid DNA from each colony and reintroducing it into fresh aliquots of *E. coli*  $\Delta metC$ . Sequencing revealed that a single genotype, encoding the ALR(Y274F) mutant, dominated the pool with a frequency of  $25/32 = 78\%$  (Table 1). All of the other selected clones also possessed the Y274F mutation, in addition to 1–2 substitutions at other positions. The ubiquity of the Y274F mutation provided strong evidence that it was essential for improving the CBL activity of ALR.

Of the selected mutations, Y274F is the only substitution located in the vicinity of the active site. Tyr-274' is found in the innermost part of the substrate entranceway (15), and it is conserved between bacterial alanine racemases (29, 30). The other nine mutated residues (Table 1) are either found on the surface of the dimeric enzyme or, in the cases of the buried residues Thr-97 and Leu-241, they appear to play simple structural roles on  $\beta$ -strands, where they have been conservatively replaced with serine and methionine in variants ALR-5 and ALR-7, respectively. Nevertheless, we were wary that distal residues can play important functional roles. Therefore, we used the Partial Order Optimum Likelihood (POOL) server (31) to predict and rank the functionally important residues in ALR, using PDB entry 2RJG as the input. For any given structure, all of the functionally important residues are typically in the top 8–10% of those ranked by POOL (31). The two Tyr-274 residues in the ALR dimer had the 11th and 16th highest POOL scores of 716 total residues (*i.e.* inside the top 2.5%). All of the other mutated residues in Table 1 had POOL scores outside of the top 35%. Thus, the additional substitutions in variants ALR-2 to ALR-8 were likely to be neutral or mildly deleterious, and we focused our attention on ALR(Y274F).

**Kinetic Characterization of ALR(Y274F)**—We purified His<sub>6</sub>-tagged ALR(Y274F) and assayed it for cystathionine  $\beta$ -lyase activity, which was found to be low but measurable (Table 2). At  $0.85 \text{ s}^{-1}\cdot\text{M}^{-1}$ , the  $k_{\text{cat}}/K_m$  was  $\sim 2.7 \times 10^6$ -fold lower than that of CBL for the same reaction. This was mainly due to a poor  $k_{\text{cat}}$  of  $\sim 10$  turnovers/h/active site. In comparison, the  $K_m$  value of ALR(Y274F) for L-cystathionine was only 85-fold lower than that of CBL for its native substrate, and at 3.3 mM, it was comparable with that of wild type ALR for L-Ala ( $K_m = 0.87 \text{ mM}$ ).

The promiscuous aminotransferase and aldolase activities of the ALR from *Geobacillus stearothermophilus* have been improved by site-directed mutagenesis (32, 33). In these examples, improvements in the new activities were accompanied by drastic reductions in the alanine racemase activity, by 3–4 orders of magnitude. In contrast, the improvement of CBL

## Reciprocal Promiscuity in Two PLP-dependent Enzymes

**TABLE 2**

**Steady state kinetic parameters of ALR, CBL, and their evolved variants**

All values are the mean  $\pm$  S.E. for three independent experiments (each including duplicate assays), except for the ALR(Y274F) assays with cystathionine, in which  $n = 6$ . The  $k_{\text{cat}}$  values are reported per active site.

Enzyme	Cystathionine $\rightarrow$ homocysteine + Pyruvate + $\text{NH}_3$			L-Ala $\rightarrow$ D-Ala		
	$k_{\text{cat}}$ $s^{-1}$	$K_m$ $mM$	$k_{\text{cat}}/K_m$ $s^{-1} \cdot M^{-1}$	$k_{\text{cat}}$ $s^{-1}$	$K_m$ $mM$	$k_{\text{cat}}/K_m$ $s^{-1} \cdot M^{-1}$
ALR	ND <sup>a</sup>	ND <sup>a</sup>	<0.85	55 $\pm$ 4	0.87 $\pm$ 0.15	6.4 $\times 10^4$
ALR(Y274F)	(2.8 $\pm$ 0.3) $\times 10^{-3}$	3.3 $\pm$ 1.0	0.85	40 $\pm$ 5	1.5 $\pm$ 0.6	2.6 $\times 10^4$
CBL	90 $\pm$ 14	(3.9 $\pm$ 0.7) $\times 10^{-2}$	2.3 $\times 10^6$	3.3 $\pm$ 0.6	51 $\pm$ 4	65
CBL(P113S)	95 $\pm$ 3	0.12 $\pm$ 0.02	7.9 $\times 10^5$	21.6 $\pm$ 0.4	58 $\pm$ 3	3.7 $\times 10^2$
CBL(P113A)	81 $\pm$ 11	0.31 $\pm$ 0.07	2.6 $\times 10^5$	4.6 $\pm$ 0.5	23 $\pm$ 5	2.0 $\times 10^2$
CBL(P113T)	30 $\pm$ 2	0.21 $\pm$ 0.02	1.4 $\times 10^5$	5.8 $\pm$ 0.5	39 $\pm$ 4	1.5 $\times 10^2$
CBL(P113Q)	96 $\pm$ 8	0.12 $\pm$ 0.02	8.0 $\times 10^5$	3.6 $\pm$ 0.6	33 $\pm$ 10	1.1 $\times 10^2$
CBL(Y111F)	0.07 $\pm$ 0.01	(2.3 $\pm$ 1.9) $\times 10^{-2}$	3.1 $\times 10^3$	ND <sup>a</sup>	ND <sup>a</sup>	$\sim$ 0.011

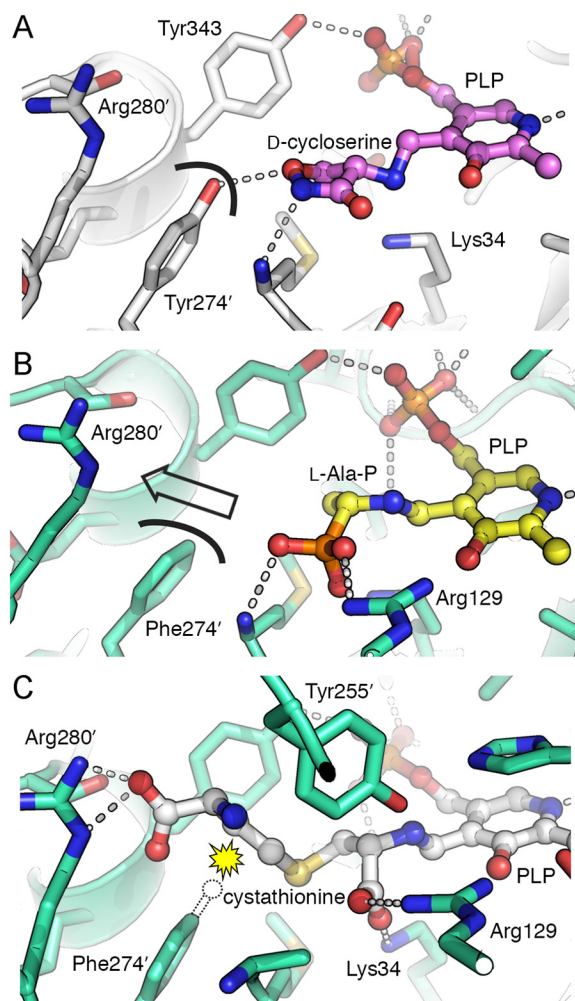
<sup>a</sup> ND means, not detected.

activity in ALR(Y274F) was accompanied by a weak trade-off in its native activity. *In vitro* alanine racemization assays showed that ALR(Y274F) had a 1.4-fold decrease in  $k_{\text{cat}}$  and a 1.7-fold increase in  $K_m$ , compared with ALR (Table 2).

**Structures of ALR(Y274F)**—Given the location of Tyr-274' at the entrance to the active site, it seemed likely that mutation to a smaller Phe residue was useful for allowing the new (and larger; Fig. 1) substrate, cystathionine, to bind productively. To confirm this hypothesis, we used x-ray crystallography to solve the structures of the His<sub>6</sub>-tagged ALR(Y274F) variant and an inhibitor-bound form. The inhibitor, (*R*)-1-aminoethylphosphonic acid (L-Ala-P), is a phosphonate analogue of L-Ala that is a substrate for the transaldimination reaction, in which the Schiff base linkage between Lys-34 and PLP in the active site of the resting enzyme is displaced by the amino group of the incoming substrate. Thus, L-Ala-P forms a stable external aldimine with PLP in the ALR active site (17).

As expected, the Y274F mutation had no significant effect on the overall architecture of ALR. A pairwise comparison of the uninhibited ALR and ALR(Y274F) dimers (from PDB entries 2RJG and 4WR3, respectively) using PDBeFold (34) gave a root-mean-square deviation of only 0.26 Å over all 718 C $\alpha$  atoms. In the wild type ALR structure, the hydroxyl group of Tyr-274' forms a hydrogen bond (2.6 Å) with a crystallographic sulfate, and it has been proposed that this sulfate mimics the position of L-Ala on its route into the active site (15). When the enzyme was inhibited by D-cycloserine, the Tyr-274' side chain hydroxyl formed a 3.2-Å hydrogen bond with the hydroxyimine moiety of the inhibitor instead (Fig. 2A). Although we also observed crystallographic sulfates in the structure of ALR(Y274F), there was neither a sulfate nor a water molecule in the position equivalent to that observed in the wild type structure. One implication is that removing the ability to form the hydrogen bond (by mutating Tyr-274 to Phe) has perturbed L-Ala binding; the increased  $K_m$  for L-Ala in ALR(Y274F) supports this hypothesis.

Binding of L-Ala-P to ALR(Y274F) did not induce any significant conformational changes in the dimer (root-mean-square deviation = 0.16 Å over 715 C $\alpha$  atoms), although we also note that the ligand-bound structure was generated by soaking. It is possible that the crystal lattice may have somewhat restricted conformational changes, although our results are consistent with those reported for the structure of the *G. stearothermophilus* ALR complexed with L-Ala-P, which was obtained by co-



**FIGURE 2. Structural analysis of the Y274F mutation in ALR.** A, in the structure of wild type ALR inhibited by D-cycloserine (PDB entry 2RJH) (15), Tyr-274' is positioned to form a hydrogen bond with the inhibitor. The curved black line shows the surface at the tyrosine hydroxyl group. B, active site of ALR(Y274F) with the L-Ala-P-PLP external aldimine, shown in the same orientation as A. The curved black line represents the surface of Phe-274'. The lack of a hydroxyl group opens a pathway from the substrate toward Arg-280' (indicated with an arrow). C, modeled position of L-cystathionine in the ALR(Y274F) active site, docked using GOLD. The distal carboxylate of the substrate is perfectly positioned to interact with Arg-280'. A dotted outline indicates the position that the Tyr-274' hydroxyl group occupies in the wild type ALR structure. The yellow starburst indicates the steric clash that would occur between this hydroxyl group and cystathionine, disfavoring substrate binding. The Y274F mutation removes this steric clash.

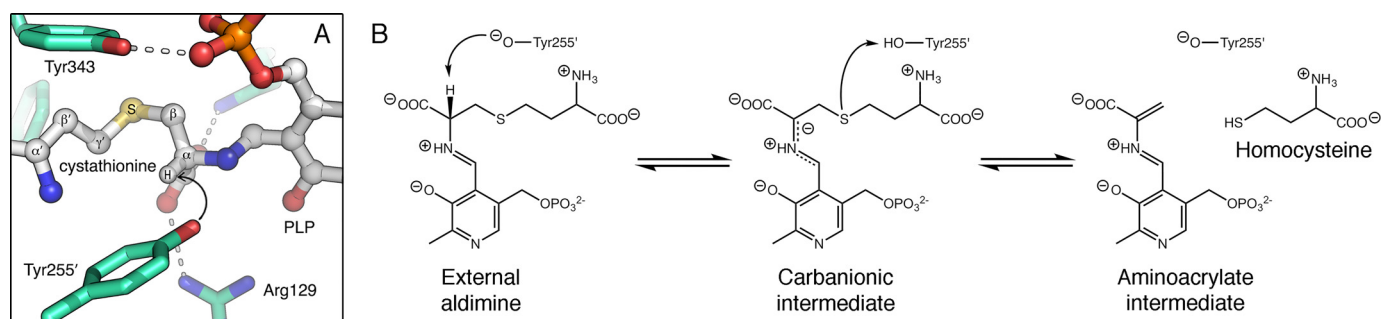


FIGURE 3. **Tyr-255' is positioned to play a catalytic role in cystathionine  $\beta$ -elimination by ALR(Y274F).** *A*, when L-cystathionine-PLP is docked into the active site using GOLD, the side chain oxygen atom of Tyr-255' is  $\sim 2.6$  Å from the proton that it is poised to abstract from the substrate (black arrow). *B*, tentative reaction scheme for ALR(Y274F)-catalyzed cystathionine  $\beta$ -elimination. Tyr-255' has an unusually low  $pK_a$  (20), so it is assumed to be in the phenolate form for proton abstraction (22). Unlike in CBL, the pyridine nitrogen is not protonated, because of the proximity of Arg-209 in the ALR active site. This results in a resonance-stabilized carbanionic intermediate of unusually high energy (45). By analogy with the CBL mechanism (35), we predict that Tyr-255' reorients to protonate the leaving group (homocysteine). This would yield the PLP derivative of aminoacrylate, which is the substrate for reverse transaldimination (35), to form iminopropionate and to regenerate the Lys-34-PLP Schiff base. The nature of the carbanionic intermediate and inefficient proton transfer from Tyr-255' to the leaving group are likely to explain the very poor turnover observed for ALR(Y274F).

crystallization (17). Formation of the external aldimine tilts the plane of the PLP pyridine ring by  $\sim 10^\circ$  toward Tyr-255'. This has been observed previously (15, 17, 21). Tyr-255' is positioned to act as the catalytic base, with a distance of 3.3 Å from its side chain hydroxyl group to  $C\alpha$  of the substrate. The side chain amino group of the catalytic acid, Lys-34, is 4.1 Å from  $C\alpha$ . As described previously (17), there was clear electron density for the L-isomer but not the D-isomer of the inhibitor, confirming that there had been no observable racemization of L-Ala-P.

The orientation of L-Ala-P in the active site of our ALR structure illuminates the role of the Y274F mutation in enhancing CBL activity. Removing the side chain hydroxyl of Tyr-274' opens a pathway toward Arg-280', without otherwise disturbing the structure of the active site (Fig. 2B). We hypothesized that cystathionine could be accommodated in the enlarged pocket formed by the Y274F mutation. To test this hypothesis, a cystathionine-PLP adduct was built and docked into the active site of ALR(Y274F) using GOLD. A scaffold matching constraint on the L-Ala-P molecule was used in the docking protocol to fix the PLP moiety of cystathionine-PLP in place, while allowing cystathionine to rotate freely and assume a biologically relevant configuration. In this docked model, the  $C\alpha$  atoms of L-Ala-P and cystathionine overlay, and the distal carboxylate of cystathionine is perfectly positioned to engage in a salt bridge with the side chain of Arg-280' (Fig. 2C). However, the presence of the wild type tyrosine at position 274' partially occludes the space between PLP and Arg-280' and therefore is predicted to result in a steric clash with cystathionine (Fig. 2C).

As with almost all PLP-dependent enzymes, ALR and CBL each catalyze transaldimination to yield an external aldimine. Furthermore, the next step in the mechanism of each enzyme is to abstract the proton from  $C\alpha$  of the substrate (21, 22, 35, 36). Although an active site lysine is responsible for this step in CBL, Tyr-255' plays the same role in ALR when the substrate is L-Ala (21). Given the shared stereochemistry of the native and promiscuous substrates (i.e. L-Ala and L-cystathionine), and the results of our docking (Fig. 2C), it is highly likely that Tyr-255' deprotonates L-cystathionine once it has formed the external aldimine in the active site of ALR(Y274F) (Fig. 3). However, further mutagenesis will be required to test this hypothesis rigorously.

*Search for Promiscuous Alanine Racemases*—With our interest in PLP-mediated promiscuity piqued, we next set out to identify *E. coli* proteins with promiscuous alanine racemase activity. We rationalized that other racemases/epimerases, other  $(\beta\alpha)_8$  barrel enzymes, or other PLP-dependent enzymes might be capable of catalyzing alanine racemization. In *E. coli*, deleting both alanine racemase genes (*alr* and *dadX*), leads to death unless D-Ala is supplied exogenously (37). Therefore, we conducted a genome-wide multicopy suppression study with *E. coli* strain MB2795 ( $\Delta alr \Delta dadX$ ). After transforming *E. coli* MB2795 with the pool of ASKA plasmids, we spread  $\sim 36,000$  transformed cells on IPTG-supplemented LB agar, in the absence of D-Ala. After 4 days, nine colonies had formed. DNA sequencing revealed that these clones harbored four plasmid-encoded ORFs as follows: *alr* (four colonies), *dadX* (two colonies), *cycA* (two colonies), and *metC* (one colony). Retransformation tests confirmed that expressing each of these ORFs led to rescue of *E. coli* MB2795 and that their growth was not a result of cross-feeding between different clones on the same agar plate.

It was expected that IPTG-induced expression of either ALR or DadX would rescue the *E. coli*  $\Delta alr \Delta dadX$  strain. *CycA* is a serine/alanine/glycine:proton symporter, the expression of which presumably improved the ability of the cells to scavenge residual D-Ala from the selective medium. Most unexpectedly, expressing the *metC* product, CBL, also appeared to overcome the D-Ala auxotrophy of *E. coli* MB2795. The implication was that CBL could catalyze alanine racemization and therefore that ALR and CBL were reciprocally promiscuous. During the course of this study, another group also independently discovered the growth of revertants in the absence of D-Ala, when CBL was up-regulated in *E. coli* BL21(DE3)  $\Delta alr \Delta dadX$  (38). Increased CBL expression in the revertants was ascribed to point mutations in the *metJ* repressor. These results are consistent with our observation that the growth of *E. coli* MB2795 was effected by overexpression of plasmid-encoded CBL.

*Improving the ALR Activity of CBL by Directed Evolution*—We probed the evolvability of CBL using epPCR. The library consisted of  $1.5 \times 10^5$  mutated *metC* variants. Sequencing the *metC* gene inserts (each 1,185 bp) from 14 randomly chosen colonies revealed a mean mutation rate of 3.0 mutations per

## Reciprocal Promiscuity in Two PLP-dependent Enzymes

**TABLE 3**

**CBL mutants selected from an epPCR library for their ability to rescue *E. coli* MB2795 in the absence of IPTG**

Variant	Frequency	Mutation(s)
CBL-1	10	P113S
CBL-2	2	P113S; 1 silent
CBL-3	2	P113A
CBL-4	1	P113A, D247E
CBL-5	1	L34M, P113A; 1 silent
CBL-6	1	P113T
CBL-7	1	P113Q; 1 silent
CBL-8	1	P113S, A237V
CBL-9	1	P113S, N321S
CBL-10	1	P113S, A325V
CBL-11	1	P113S, F331T
CBL-12	1	P113S, A246V; 1 silent

clone, corresponding to 2.0 amino acid substitutions per CBL variant (28). The cloned library was introduced into *E. coli* MB2795 and spread on LB agar with 5  $\mu$ M IPTG. In the absence of exogenous D-Ala, the parental clone (*E. coli* MB2795 expressing His<sub>6</sub>-tagged CBL) took 5–7 days to form colonies under these conditions. In contrast, ~4% of the library had formed colonies after incubation for 42 h. Previously, we observed some leaky expression from the T5-*lacO* promoter of pCA24N (26). Therefore, we retested 48 of the fast-growing clones on LB agar with no IPTG, and we identified 23 that formed colonies after 24 h of incubation.

DNA sequencing revealed that the population of 23 fast-growing clones was made up of 12 different variants (Table 3). Notably, all variants contained mutations at CBL residue Pro-113, immediately suggesting that this amino acid was a critical determinant of alanine racemization activity. Pro-113 was mutated to either serine, alanine, threonine, or glutamine. The most common mutation was P113S, which appeared in 17 of the 23 clones. This was followed by the P113A mutation, which was found in four clones. Single clones (CBL-6 and CBL-7) carried the P113T and P113Q mutations. Mutations were also identified at seven additional positions, in single variants (Table 3).

None of the residues that were mutated in CBL are known to possess direct catalytic or substrate-binding roles. One residue (Leu-34, mutated to methionine in variant CBL-5) is located on the monomer-monomer interface of the N-terminal tetramerization domain (11). Four residues (Pro-113, Ala-237, Ala-246, and Asp-247) are in the middle domain of CBL, where the active site and PLP are located, but none contact the substrate or cofactor directly. The three remaining residues (Asn-321, Ala-325, and Phe-331) are in the C-terminal domain and are located on a solvent-exposed helix. Given the predominance of mutations at Pro-113, we surmised that the additional mutations were likely to be either neutral or mildly deleterious, with respect to the alanine racemization activity of CBL. Therefore, subsequent characterization focused on variants CBL-1, CBL-3, CBL-6, and CBL-7 (Table 3).

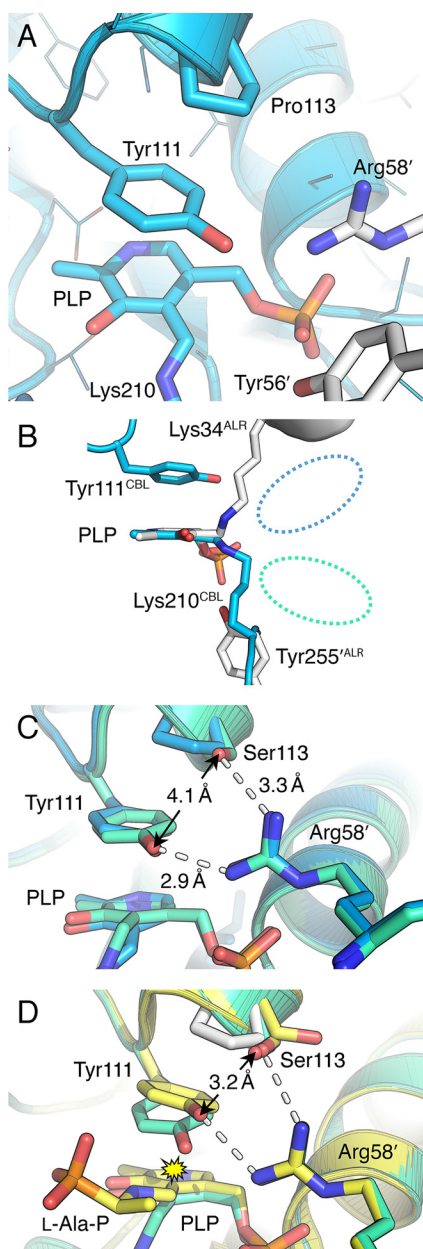
**Kinetic Characterization of CBL Variants**—We expressed and purified CBL and the four Pro-113 variants. Their L-alanine racemization activities were readily detected by using a coupled assay with D-amino acid oxidase (for converting the product, D-Ala, to pyruvate) and lactate dehydrogenase (for the NADH-linked reduction of pyruvate to lactate). We excluded the pos-

sibility that CBL might be acting as a deaminase (*i.e.* converting L-Ala into pyruvate), because a positive signal was only observed when D-amino acid oxidase was present. This confirmed that D-Ala was produced from L-Ala by CBL.

At 65 s<sup>-1</sup>·M<sup>-1</sup>, the  $k_{\text{cat}}/K_m$  of CBL for alanine racemization was almost 1,000-fold lower than ALR itself (Table 2). This was due to CBL possessing both a substantially larger  $K_m$  for L-Ala (51 mM) and also a 17-fold lower  $k_{\text{cat}}$  (3.3 s<sup>-1</sup>). The evolved CBL variants had overall improvements in catalytic efficiencies that ranged from 1.7-fold for CBL(P113Q) to 5.7-fold for CBL(P113S). In the cases of the CBL P113A, P113T, and P113Q variants, these increases were mainly brought about by decreases in the  $K_m$  value for L-Ala. In contrast, CBL(P113S) showed a 6.5-fold increase in  $k_{\text{cat}}$ , whereas its  $K_m$  value for L-Ala was similar to CBL (Table 2). Consistent with its predominance in the library selection experiment (Table 3), CBL(P113S) was the most active variant overall, with a  $k_{\text{cat}}/K_m$  that was only 170-fold lower than ALR.

The cystathionine  $\beta$ -elimination activities of the CBL Pro-113 variants were also assayed, to gauge the trade-off between the native and evolving activities. The variants showed cystathionine  $\beta$ -elimination activities that were reduced by 3–16-fold, compared with CBL itself (Table 2). CBL(P113T) was the only variant with a significantly reduced  $k_{\text{cat}}$ ; however, all four variants had increased  $K_m$  values for L-cystathionine. In the P113A, P113T, and P113Q CBL variants, the mutations decreased cystathionine  $\beta$ -elimination activity by more than they increased alanine racemization activity. In contrast, CBL(P113S) showed a 2.9-fold reduction in cystathionine  $\beta$ -elimination, accompanied by a 5.7-fold increase in alanine racemization.

**Toward a Functional Role for the Pro-113 Variants**—Pro-113 is close to the active site of CBL, being located at the N-terminal end of helix 6 and packing directly above the active site residue Tyr-111 (Fig. 4A). Inspection of the wild type CBL structure (11) immediately suggested Tyr-111 and Lys-210 as candidates to act as the catalytic acid and base for alanine racemization by CBL, albeit positioned on the opposite faces of PLP than Tyr-255' and Lys-34 in ALR (Fig. 4B). This orientation would mean that Lys-210 was responsible for deprotonating L-Ala (instead of Tyr-255' in ALR) and that Tyr-111 was protonating the planar intermediate to yield D-Ala. We probed the role of Tyr-111 by constructing the site-directed mutant CBL(Y111F). A previous analysis of this variant showed that it was prone to inhibition by DTNB, while clarifying that Tyr-111 does not play a catalytic role in proton transfer during the CBL reaction (39). Consistent with these results, we observed a steep drop-off in  $k_{\text{cat}}$ , but not  $K_m$ , for CBL(Y111F) in our DTNB-coupled assays for cystathionine  $\beta$ -elimination (Table 2). Of more relevance here, and as predicted, CBL(Y111F) had ALR activity that was barely detectable above baseline. Although we could not saturate the enzyme, we estimated its  $k_{\text{cat}}/K_m$  to be ~0.011 s<sup>-1</sup>·M<sup>-1</sup>, which is almost 6,000-fold lower than the ALR activity of wild type CBL (Table 2). This confirmed a critical catalytic role for Tyr-111 in alanine racemization and led us to hypothesize that the mutations at Pro-113 had been selected because they reoriented Tyr-111 to act more effectively as the catalytic acid in the conversion of L-Ala to D-Ala.



**FIGURE 4. Structural analysis of the P113S mutation in CBL.** *A*, in wild type CBL (PDB entry 1CL1) Pro-113 is positioned above Tyr-111, which in turn makes a stacking interaction with the pyridine ring of PLP (11). *B*, overlay of the catalytic residues in CBL (blue) and ALR (white), with the structures oriented so that their PLP cofactors superimpose. The approximate position occupied by incoming substrate is indicated with a dashed blue oval for CBL and a dashed green oval for ALR. *C*, CBL(P113S) (green) superimposed upon CBL (blue). Although there are no significant changes in the positions of active site residues, Arg-58' is now within hydrogen bonding distance of Ser-113 in the mutant. *D*, overlay of the unliganded CBL(P113S) structure (green) and CBL(P113S) with an L-Ala-P-PLP external aldimine (yellow). Wild type residue Pro-113 is also shown in white to highlight the effect of the P113S substitution. Two rotamers of Ser-113 were observed in the inhibited structure; both are shown. The two hydrogen bonds involving Arg-58' are the same length as in *C*. The yellow starburst highlights the steric clash that would occur if Tyr-111 did not pivot away from PLP in the inhibited structure.

**Structures of CBL(P113S)**—To test our hypothesis, we solved the unliganded and L-Ala-P-bound structures of CBL(P113S). To our surprise, the mutation did not significantly affect the position of Tyr-111 in the unliganded structure (Fig. 4C). As observed in wild type CBL, the Tyr-111 side chain forms a  $\pi$

stacking interaction with the pyridine ring of PLP and a hydrogen bond (2.9 Å) with the side chain of Arg-58' from an adjacent monomer (*i.e.* Arg-58'). This hydrogen bond is the same length in the wild type structure (11). In CBL, Arg-58' also makes a salt bridge with the phosphate group of PLP (11). All of these interactions remain in the CBL(P113S) structure. In the mutant, nitrogen atom NH1 of Arg-58' is now within hydrogen bonding distance of the Ser-113 side chain (N-O distance of 3.3 Å; Fig. 4C).

We readily obtained a CBL(P113S) structure in which the ALR inhibitor, L-Ala-P, had entered the active site and formed a stable external aldimine. As with ALR(Y274F), this was obtained by soaking a pre-formed crystal with the inhibitor. Also, as we observed with ALR(Y274F), the inhibitor was in the L-isomer, and formation of the external aldimine was accompanied by a tilt of PLP by  $\sim 10^\circ$  away from Lys-210 (Fig. 4D). This induces a tilt in the Tyr-111 side chain by the same amount. These changes are smaller than those reported for the binding of a cystathionine analogue, L-aminoethoxyvinylglycine (AVG) to the wild type enzyme, which induced tilts of  $\sim 30^\circ$  in PLP and Tyr-111 (35). More substantially, the side chain also pivots by  $\sim 25^\circ$  along a perpendicular axis; therefore, the overall movement of Tyr-111 is best described as “up and away” from the L-Ala-P-PLP adduct (Fig. 4D). This repositioning of Tyr-111 avoids potential steric clashes between its hydroxyl group and C4' of the cofactor ( $\sim 2.4$  Å), as well as with C $\beta$  of the L-Ala side chain ( $\sim 2.6$  Å). By pivoting away, these “clash” distances increase to  $\sim 3.8$  and  $\sim 3.6$  Å, respectively (Fig. 4D).

The repositioning of Tyr-111 also brings its hydroxyl group closer to the side chain of Ser-113. In the inhibited structure, the electron density for Ser-113 is consistent with the side chain adopting two rotamers, with approximately equal occupancy (Fig. 4D). In one of these, the side chain is rotated toward bulk solvent. In the other, it occupies the same position as it does in the unliganded structure. In this latter conformation, the distance between the hydroxyl groups of Tyr-111 and Ser-113 decreases from 4.1 Å in the unliganded structure (Fig. 4C) to 3.2 Å when L-Ala-P is bound (Fig. 4D). Arg-58' does not alter its position, and the 2.9-Å hydrogen bond with Tyr-111 remains intact.

We anticipated that the movement of Tyr-111 toward Ser-113 might alter the  $pK_a$  of its hydroxyl group and therefore its ability to act as the catalytic acid during L-Ala racemization. The relevant side chain  $pK_a$  values were estimated using PROPKA 3.1 (40). This analysis (Table 4) suggested that the introduction of the P113S mutation lowered the side chain  $pK_a$  values of both Tyr-111 ( $\Delta pK_a = -3.2$ ) and Arg-58' ( $\Delta pK_a = -5.5$ ). Binding of AVG to wild type CBL increases the  $pK_a$  of Tyr-111 by over 2 units, although the  $pK_a$  of Arg-58' remains unchanged. In contrast, the binding of L-Ala-P to CBL(P113S) does not appear to alter the  $pK_a$  of Tyr-111, provided Ser-113 is oriented toward it. When Ser-113 is oriented toward solvent instead, the  $pK_a$  of Tyr-111 increases 3 units on binding the inhibitor. In their inhibitor-bound external aldimine forms, with Ser-113 oriented toward the active site, the difference in theoretical Tyr-111  $pK_a$  values between CBL and CBL(P113S) is estimated to be 5.6 units.

## Reciprocal Promiscuity in Two PLP-dependent Enzymes

**TABLE 4**

Selected side chain  $pK_a$  values in liganded and unliganded CBL variants, as estimated by PROPKA 3.1

Enzyme and ligand (Ser-113 rotamer)	$pK_a$ of Tyr-111	$pK_a$ of Arg-58'	PROPKA input (PDB code)
CBL	14.7	22.4	1CL1
CBL(P113S)	11.5	16.9	4ITG
CBL; AVG	17.0	22.5	1CL2
CBL(P113S); L-Ala-P (Ser-113 to Tyr-111)	11.4	19.8	4ITX
CBL(P113S); L-Ala-P (Ser-113 to solvent)	14.6	19.9	4ITX

### Discussion

*Reciprocally Promiscuous but Non-homologous Enzymes*—The prevailing model for enzyme evolution invokes primordial cells containing a small number of genes, with each of these genes encoding a multifunctional protein (41). From this rudimentary starting point, gene duplication and divergence could give rise to the highly active, highly specific enzymes that are hallmarks of modern metabolic pathways. In many cases, the promiscuous activities of modern specialist enzymes are assumed to be evolutionary vestiges of their ancient multifunctional ancestors. Indeed, patterns of shared and overlapping promiscuous activities are common among the homologous members of enzyme superfamilies (42–44).

To our knowledge, this study reports the first example of reciprocal promiscuity between non-homologous enzymes. Instead of sequence similarity or a shared fold, ALR and CBL each utilize the powerful catalytic cofactor PLP.

*ALR Is a Promiscuous and Evolvable CBL*—The PLP-mediated cystathionine  $\beta$ -elimination activity of ALR was sufficient for enabling the slow growth of *E. coli*  $\Delta metC$  cells. A single round of directed evolution yielded ALR(Y274F), which was amenable both to kinetic characterization and to structure determination. The conservative Y274F mutation enlarged the active site pocket just enough that cystathionine could enter it and form an external aldimine with PLP (Fig. 2).

Although ALR(Y274F) is able to bind cystathionine productively, its turnover number for  $\beta$ -elimination is very low ( $k_{cat} = 0.0028 \text{ s}^{-1} \approx 10 \text{ h}^{-1}$ ). In CBL, Lys-210 mediates proton transfer from  $C\alpha$  to  $S\gamma$ , with the elimination of homocysteine being rate-determining (36). In ALR, Tyr-255' deprotonates L-Ala during the racemization reaction, and our docking analysis (Fig. 3A) shows that it is similarly well placed to deprotonate L-cystathionine. However, it is less clear whether Tyr-255' is able to reorient to transfer this proton to the homocysteine leaving group (Fig. 3B). Furthermore, and in contrast to most PLP-dependent enzymes, ALR deprotonates the cofactor by placing an arginine (Arg-209 in the *E. coli* enzyme) adjacent to the nitrogen of the pyridine ring (4). This results in the formation of an unusually unstable carbanionic quinonoid intermediate during alanine racemization (45). In turn, this favors rapid reprotonation of  $C\alpha$  (i.e. racemization) by either Lys-34 or Tyr-255', instead of the (slower) reorientation of these residues to protonate  $C4'$  of the cofactor, which would lead to transamination (4). In CBL, the presence of Asp-185 ensures that the pyridine nitrogen is protonated (11), the quinonoid is stabilized, and Lys-210 can reorient to mediate proton transfer. In combination, the short-lived carbanionic intermediate and inefficient

proton transfer to  $S\gamma$  appear to ensure that ALR(Y274F) is a rudimentary CBL, at best.

*CBL Is a Promiscuous and Evolvable ALR*—Having characterized the promiscuous activity of ALR, we were surprised to discover that *E. coli* CBL is also a promiscuous ALR. As with ALR, the promiscuous activity of CBL could also be readily improved in a single round of epPCR and selection. The promiscuous activity of ALR(Y274F) was characterized by a good Michaelis constant but poor turnover. In contrast, the promiscuous activities of CBL and its Pro-113 mutants were characterized by relatively poor Michaelis constants but good turnover numbers (Table 2). Most notably, the CBL(P113S) variant had a  $k_{cat}$  for L-Ala racemization that was  $22 \text{ s}^{-1}$  per active site. This is comparable with the turnover number of ALR for the same reaction ( $k_{cat} = 55 \text{ s}^{-1}$  per active site).

In the CBL active site, Lys-210 is ideally positioned to act as the catalytic base that deprotonates L-Ala after formation of the external aldimine (Fig. 5). It plays the same role in deprotonating the native substrate, L-cystathionine (36). Furthermore, Lys-210 is tethered in place by hydrogen bonds with the side chain hydroxyl groups of Tyr-56' and Ser-339. This constrains the reaction specificity of CBL, by disfavoring protonation at the  $C4'$  position of the cofactor (39). As discussed above, ALR achieves the same outcome by destabilizing the quinonoid intermediate. Each enzyme has evolved to minimize competing transamination reactions, coincidentally without narrowing its reaction specificity with respect to racemization and  $\beta$ -elimination.

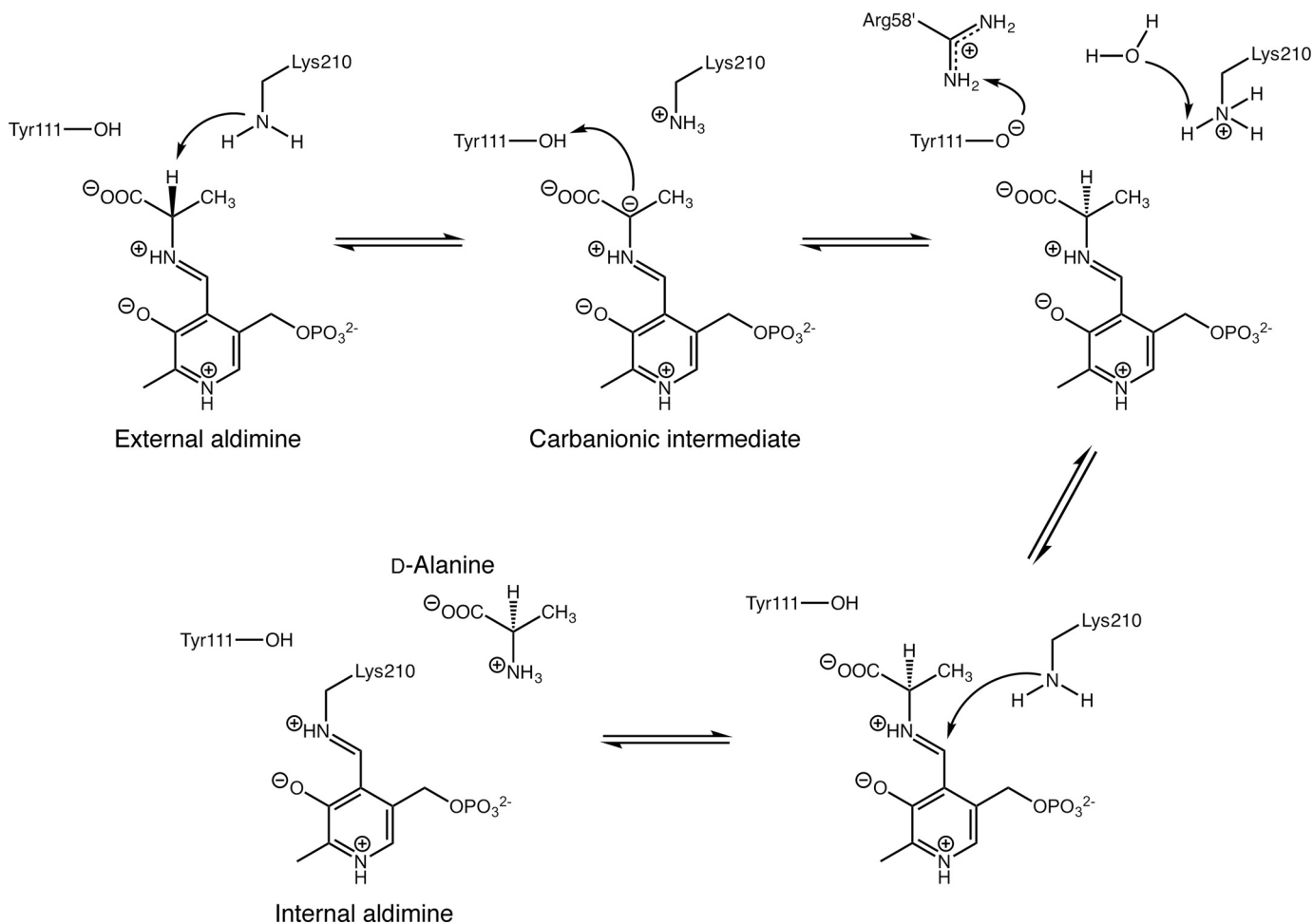
A second parallel between the two active sites concerns the roles of Tyr-56' in CBL and Tyr-343 in ALR. Both tyrosines help to anchor the cofactor phosphate, via hydrogen bonds of 2.7 and 2.6 Å, respectively. Each is also oriented so that, during the racemization reaction, the side chain of the substrate must move toward and past it to adopt the planar carbanionic intermediate and then the racemized product. In *G. stearothermophilus* ALR, mutation of the equivalent tyrosine (Tyr-354) to Asn yielded a bi-functional alanine/serine racemase (46). It remains to be tested whether a comparable mutation could broaden the substrate range of CBL acting as a racemase.

Completing the alanine racemization reaction requires protonation of the carbanionic intermediate from the opposite face than Lys-210. The structural data and our analysis of the CBL(Y111F) variant imply that Tyr-111 is the catalytic acid for racemization of L- to D-alanine (Fig. 5). An equivalent mutation (Y265F) has been made in the *G. stearothermophilus* ALR (19). In this enzyme, Tyr-265' (Tyr-255' in *E. coli* ALR) is the catalytic acid for the racemization of D- to L-alanine. We observed a 6,000-fold decrease in L-Ala racemization activity (from a low starting point;  $k_{cat}/K_m = 65 \text{ s}^{-1} \cdot \text{M}^{-1}$ ) for CBL(Y111F), whereas the D-Ala racemization activity of *G. stearothermophilus* ALR(Y265F) was decreased by 60,000-fold (19).

In our L-Ala-P-PLP structure (Fig. 4D), the distance from the exocyclic oxygen atom of Tyr-111 to  $C\alpha$  of the substrate analogue is 4.2 Å. Attaining the planar intermediate will bring  $C\alpha$  closer to the catalytic acid, albeit in an orientation that is still not optimal for proton transfer. Nevertheless, the equivalent distance in our L-Ala-P-PLP structure of ALR(Y274F), between  $C\alpha$  and the catalytic acid Lys-34, is almost identical at 4.1 Å.



## Reciprocal Promiscuity in Two PLP-dependent Enzymes



**FIGURE 5. Reaction scheme for racemization of L-alanine, catalyzed by CBL(P113S).** Estimates from PROPKA suggest that Tyr-111 is likely to be protonated when the L-Ala-PLP external aldimine is formed. When CBL catalyzes cystathionine  $\beta$ -elimination, the carbanionic intermediate is stabilized as the ketamine quinonoid (36). Here, we show instead the resonance structure of the carbanionic intermediate that is consistent with the ALR mechanism (22). Also, as done previously (22), we have shown the proton exchanges at Tyr-111 (involving Arg-58') and Lys-210 occurring from the same intermediate, but this is not a necessity. Reverse transaldimination to regenerate the Lys-210-PLP Schiff base is assumed to occur as described for the cystathionine  $\beta$ -elimination reaction.

Previous structural and mutagenesis studies have interrogated the role of Tyr-111 in the native activity of CBL. Early structural work suggested that it might act as a catalytic acid/base, abstracting a proton from the  $\alpha$ -amino group of cystathionine during transaldimination and then donating it to the aminoacrylate that is formed after homocysteine has been eliminated (11). This model required the nearby side chain of Arg-58' to lower the  $pK_a$  of the Tyr-111 side chain to the extent that it was a phenolate ion when substrate entered the active site. However, detailed kinetic analyses have ruled out a role for Tyr-111 in proton transfer (39). Our data on the alanine racemase activities of CBL and CBL(P113S) are consistent with the latter study. Our estimate of the Tyr-111 side chain  $pK_a$  (14.7; Table 4) suggests that it will be uncharged in the substrate-free internal aldimine form of the enzyme. As predicted previously (11), this  $pK_a$  increases when the substrate analogue, AVG, enters and increases the hydrophobicity of the active site. However, PROPKA 3.1 estimates that the presence of serine at position 113 will lower the theoretical  $pK_a$  of Tyr-111 to  $\sim 11.5$ , with or without bound substrate. This is likely due to the new hydrogen bond between Ser-113 and Arg-58' (Fig. 4C) lowering the  $pK_a$  of the latter and establishing a proton relay. The implica-

tion is that Tyr-111 is protonated when L-Ala (or L-Ala-P) binds but is better able to donate that proton during racemization (Fig. 5). By sampling two side chain conformations (Fig. 4D), Ser-113 also provides a route to shuttle a proton from bulk solvent back to Tyr-111, via Arg-58', to complete the catalytic cycle (Fig. 5). Directed evolution acted to optimize the side chain  $pK_a$  of a key catalytic residue. This was a less intuitive outcome than our initial prediction (*i.e.* repositioning of Tyr-111 to effect protonation of the carbanionic intermediate), and it highlights the ability of directed evolution to yield mechanistically unexpected solutions to challenges in enzyme engineering and evolution.

The native activity of CBL was reduced in all four of the Pro-113 variants that we characterized, by 3–16-fold (Table 2). In each mutant, most of this reduction in  $k_{cat}/K_m$  resulted from an increase in the  $K_m$  value for cystathionine. One possible explanation involves Arg-58', which plays an additional role by interacting with the distal carboxylate of cystathionine (47). Compared with proline, the side chains that we introduced at position 113 are each likely to partially occlude the space for this interaction, and/or alter the bonding capability of Arg-58' (such as by introducing a new hydrogen bond in the case of

## Reciprocal Promiscuity in Two PLP-dependent Enzymes

Ser-113; Fig. 4C). Further structural studies with cystathionine analogues, such as AVG or the bis-hydrazino molecules described recently (48), will be useful for clarifying the effects on CBL activity of mutating Pro-113.

**Evolutionary Considerations**—PLP-dependent enzymes have attracted considerable interest as exemplars of protein evolution, due to the impressive diversity of their folds and functions. A parsimonious model suggests that the initial step in the evolution of each fold type was to constrain the reactivity of PLP to an overall reaction type, with narrowed substrate specificities arising later (49). This model would predict, for example, that all cystathionine  $\beta$ -lyases belong to fold type I and all alanine racemases belong to fold type III. However, our study adds to the body of evidence that convergent evolution has been common (2, 49, 50) and that there can be very short mutational trajectories between PLP-dependent enzymes with different reaction specificities.

Enzymes that catalyze  $\beta$ -elimination reactions and adopt fold types I, II, and IV have been characterized (2). It has been noted that this example of convergent evolution is unsurprising, because  $\beta$ -elimination of a good leaving group is highly facile and may simply require mutations that improve the binding of substrates with  $\beta$  leaving groups (50). Our discovery that ALR is a promiscuous CBL adds fold type III to the list of scaffolds that can support  $\beta$ -elimination reactions. As predicted (50), the first step in evolving a more efficient enzyme, the Y274F mutation, was to improve substrate binding. Further evolutionary enhancement of catalytic efficiency is likely to involve more wholesale changes to the active site. An ALR-based scaffold with native levels of CBL activity is likely to require the R209E mutation (to protonate the pyridine ring of the cofactor) and perhaps the mutation of Tyr-255' to Lys to introduce the conformational flexibility required for effecting efficient proton transfer from C $\alpha$  to S $\gamma$ . It is noteworthy that the former mutation improved the transamination activity of *G. stearothermophilus* ALR by 5.4-fold (32).

Although the *in vitro* kinetic parameters of ALR(Y274F) are poor, it is important to note that selection acts to improve organismal fitness (and not necessarily  $k_{\text{cat}}$  or  $K_m$ ). The intracellular performance of ALR(Y274F) is substantially better than its kinetics suggest. Although its  $k_{\text{cat}}/K_m$  value for the CBL reaction is  $2.7 \times 10^6$ -fold lower than CBL itself, cells expressing ALR(Y274F) formed regular colonies in <48 h (~2-fold slower than the same strain expressing CBL). One factor contributing to this apparent discrepancy between *in vitro* activity and *in vivo* performance is likely to be the high enzyme concentrations found *in vivo*. The conditions for inducing ALR(Y274F) expression were relatively mild (5  $\mu\text{M}$  IPTG). Nevertheless, the volume of the *E. coli* cytoplasm is very small (~1.25 fl), and in a previous study we estimated that the concentration of the rescuing enzyme may approach the millimolar range (26). We recently catalogued a number of comparable examples from other groups, in which enzymes have been replaced by weakly active alternatives, but cell growth has barely been diminished (51).

Conversely, another implication may be that CBL is “better than it needs to be,” *i.e.* mutations that erode its kinetic parameters for cystathionine  $\beta$ -elimination are unlikely to affect

growth of the host cell. There is mounting evidence that many enzymes share this feature, presumably reflecting past evolutionary pressures (51).

The promiscuous activity of CBL was significantly greater than that of ALR and confirms that fold type I is compatible with racemase activity. Indeed, the alanine racemase from the fungus *Tolypocladium inflatum*, which provides the D-Ala required for cyclosporin A biosynthesis, is predicted to adopt this fold (52). With a  $k_{\text{cat}}$  of  $3.8 \text{ s}^{-1}$  and a  $k_{\text{cat}}/K_m$  of  $5.4 \times 10^2 \text{ s}^{-1}\cdot\text{M}^{-1}$ , the *T. inflatum* enzyme has similar ALR activity to CBL(P113S) (Table 2).

CBL(P113S) has a  $k_{\text{cat}}/K_m$  value for L-Ala racemization that is only 170-fold lower than ALR itself. Thus, it mimics a bi-functional ancestor of the sort envisaged by Jensen (41). Furthermore, genome reduction is a pervasive mode of bacterial evolution (53). The evolutionary emergence of a bi-functional enzyme such as CBL(P113S) could remove any need to retain a specialized ALR and allow a bacterium to streamline its genome by losing its *alr* gene. A comparable event has occurred in the evolution of *Chlamydia trachomatis*, which has a small genome (1.0 Mbp) and uses a bi-functional TrpF enzyme to catalyze reactions in the tryptophan and tetrahydrofolate biosynthetic pathways (54). Another bacterium with a small genome (2.8 Mbp), *Treponema denticola*, uses a homologue of CBL, cystalysin, to catalyze cleavage of L-cysteine to pyruvate,  $\text{NH}_3$ , and  $\text{H}_2\text{S}$ . This enzyme has promiscuous ALR activity (55), although *T. denticola* also contains a specialist alanine racemase (with NCBI accession number NP\_971703). One might predict that *T. denticola* will evolve to lose the gene for this racemase in future. This appears to have occurred already in *Wolbachia* strains (with genomes of ~1.1 Mbp), in which it has been reported that CBL has ALR activity (albeit not quantified) (56). Although *Wolbachia* strains lack a specialist alanine racemase, they are also methionine auxotrophs (57), so CBL may be maintained solely for its ALR functionality in these organisms.

**Concluding Remarks**—We have demonstrated that two non-homologous specialized enzymes from *E. coli* are as little as one mutation away from being physiologically relevant replacements for each other. This emphasizes the catalytic versatility of PLP and the ease with which new functions can evolve on different PLP-binding protein scaffolds.

With the cofactor driving catalysis, there appears to be nothing that inherently constrains one function to one fold among the PLP-dependent enzymes. Given the preponderance of cofactor-mediated promiscuity, the short mutational routes between functions, historical contingency, and the sheer depth of evolutionary time, this implies that most (if not all) PLP-dependent functions will ultimately be discovered on most (if not all) types of PLP-binding fold. It will be (and already will have been) difficult to assign physiological functions correctly to PLP-binding enzyme from genome data alone. Our data also demonstrate that overlapping patterns of catalytic promiscuity do not necessarily imply homology in PLP-dependent enzymes. Thus, the evolutionary processes that have given rise to this diverse group of enzymes are much richer than those encapsulated by simple models that invoke tidy progressions from multifunctional ancestors to specialist descendants.

## Experimental Procedures

**Materials**—Molecular biology enzymes were from New England Biolabs (Ipswich, MA). Oligonucleotides were from Integrated DNA Technologies (Coralville, IA). Chemicals were from Sigma, unless noted otherwise. IPTG was from Gold Bio-technology (St. Louis, MO). L-Cystathionine was dissolved in 0.04 N HCl to make 25 mM stocks. PLP was dissolved in 1 N HCl to make 100 mM stocks. DNA sequencing was done by Macrogen (South Korea).

**Bacterial Strains and Media**—The *E. coli*  $\Delta metC$  and  $\Delta metE$  strains were from the Keio collection (58). Under non-selective conditions, these strains were maintained in LB. *E. coli* strain MB2795 ( $\Delta alr \Delta dadX$ ) was grown in the presence of 0.5 mM D-Ala under non-selective conditions. All plasmids were maintained using appropriate antibiotics as follows: carbenicillin, 50  $\mu\text{g}\cdot\text{ml}^{-1}$ ; chloramphenicol, 34  $\mu\text{g}\cdot\text{ml}^{-1}$ ; or kanamycin, 30  $\mu\text{g}\cdot\text{ml}^{-1}$ . M9/glucose was used as the selective growth medium for the *E. coli*  $\Delta metC$  and  $\Delta metE$  strains. M9 plates contained 1  $\times$  M9 salts, 2 mM  $\text{MgSO}_4$ , 0.1 mM  $\text{CaCl}_2$ , 0.4% (w/v) glucose, and 1.5% (w/v) agar. For prolonged incubations (>2 days), all agar plates were placed in airtight plastic containers lined with moistened paper towels to prevent overdrying of the agar.

**Selection of Suppressors**—The pooled plasmids of the ASKA ORF library (23) were used to transform *E. coli* MB2795 by electroporation. Recovered cells were washed three times with 1  $\times$  M9 salts before an aliquot containing  $\sim 36,000$  transformed cells (estimated by plating dilutions on non-selective medium) was spread on LB-chloramphenicol supplemented with 50  $\mu\text{M}$  IPTG, but in the absence of D-Ala. The plate was incubated at 30 °C for 7 days, and colonies that formed were used as the source of template DNA for PCR, using the backbone-specific primers pCA24N.for and pCA24N.rev (10). The resulting PCR products were sequenced to reveal the identity of the ORFs.

**Retransformation Tests**—After each rescuing ORF had been identified, the plasmid from the corresponding ASKA clone was isolated. Aliquots of *E. coli* MB2795 were transformed with each purified plasmid. After recovering the cells in SOC medium supplemented with D-Ala at 37 °C for 1 h, the cells were washed with 1  $\times$  M9 medium. A total of 250–1,000 transformed cells were spread on LB-chloramphenicol plates containing 50  $\mu\text{M}$  IPTG. Each agar plate was incubated at 30 °C for at least 7 days.

**Growth Complementation Assays**—For complementation of the *E. coli*  $\Delta metC$  and  $\Delta metE$  strains, an overnight culture of each test clone was centrifuged at 12,000  $\times g$  for 4 min. The cell pellet was washed three times with 1 ml of 1  $\times$  M9 medium. The final cell resuspension was diluted to an  $A_{600}$  of 0.1 using 1  $\times$  M9 medium. A 2.5- $\mu\text{l}$  aliquot was spread on M9/glucose plates containing 50  $\mu\text{M}$  IPTG. Cell counts were obtained by plating appropriate dilutions of each culture on LB agar plates. The M9/glucose plates were incubated at 28 °C for at least 7 days, whereas the LB agar plates were incubated overnight at 37 °C. Complementation tests with *E. coli* MB2795 were carried out the same way, except that LB was used to wash the cell pellet, growth tests were on LB agar plates with IPTG (50  $\mu\text{M}$ ), and total cell counts were obtained by spreading dilutions of the

cell resuspension on LB-chloramphenicol plates with 0.5 mM D-Ala.

**Enzyme Expression and Purification**—All enzyme variants were expressed with His<sub>6</sub> tags at their N termini, from the respective ASKA plasmids (with their C-terminal GFP tags removed). A 4-ml culture of *E. coli* cells harboring each expression vector was grown to saturation, before being used to inoculate 250 ml of LB broth containing appropriate antibiotics. At  $A_{600} \sim 0.5$ , IPTG (0.5 mM, final concentration) was added to the culture. Protein overexpression was carried out at 28 °C for another 5–16 h. Cells were then collected by centrifuging at 5,000  $\times g$  and 4 °C for 20 min.

Column Buffer (CB) consisted of 50 mM potassium phosphate buffer (pH 8.0), 300 mM NaCl, 3 mM imidazole, 1 mM  $\beta$ -mercaptoethanol, 10% (v/v) glycerol, and 10  $\mu\text{M}$  PLP. Each cell pellet was resuspended with 4 ml of this buffer/g (wet pellet weight). Lysozyme (0.5  $\text{mg}\cdot\text{ml}^{-1}$ ), DNase I (2 units), and protease inhibitor mixture (50  $\mu\text{l}$ ) were added, and the suspension was incubated on ice for 45 min. Cell lysis was by 10 cycles of sonication (amplitude 45%; ultrasonic processor S-4000, Misonix) on ice. The lysate was clarified by centrifugation at 20,000  $\times g$  for 35 min at 4 °C. The soluble fraction was incubated with 1 ml of Talon<sup>®</sup> metal affinity resin (Clontech) at 4 °C for 2 h. Next, the resin was washed four times with 3 ml of CB, and finally, one time with 3 ml of CB that contained 15 mM imidazole. The His<sub>6</sub>-tagged enzyme was eluted with 3 ml of Elution Buffer (50 mM potassium phosphate buffer (pH 8.0), 300 mM NaCl, 150 mM imidazole, 1 mM  $\beta$ -mercaptoethanol, 10% (v/v) glycerol). All enzyme variants were judged to be >95% pure by SDS-PAGE analysis.

Purified enzymes were dialyzed extensively against Dialysis Buffer (50 mM potassium phosphate (pH 8.0), 150 mM NaCl, 10% (v/v) glycerol) at 4 °C before being concentrated using a centrifugal filtration concentrator (GE Healthcare, Little Chalfont, UK). Aliquots of each purified enzyme were stored at either 4 or  $-80$  °C. Total enzyme concentrations were quantified using molar extinction coefficients that were calculated as described previously (59). The concentrations of PLP-bound enzymes were determined using a molar extinction coefficient of 8,450  $\text{M}^{-1}\cdot\text{cm}^{-1}$  for PLP (60). In all batches of purified ALR, ALR(Y274F), and CBL, at least 70% of the total protein was PLP-bound and hence catalytically active. For the CBL variants, the proportions of PLP-bound protein were 30–50%. All enzymes were used within 2 weeks of being purified.

**Steady State Kinetic Assays**—Activity assays were performed using a Cary 100 UV-visible spectrophotometer (Agilent Technologies; Santa Clara, CA). Reactions were measured in quartz cuvettes with a path length of 1 cm.

A continuous coupled assay, based on that described previously (61), was used for detecting the racemization of L-Ala to D-Ala. The 1-ml reaction mixtures contained 100 mM CHES (pH 9.0), 0.2 mM NADH, 1 unit of D-amino acid oxidase, and 120 units of lactate dehydrogenase, with the L-Ala concentration varied. Each mixture was pre-incubated at 37 °C for 2 min, and then enzyme was added to initiate the reaction. The final enzyme concentrations in each reaction were as follows: ALR, 8.4 nM; and CBL variants, 29–158 nM. The oxidation of NADH to  $\text{NAD}^+$  during the reaction was monitored by a decrease in

## Reciprocal Promiscuity in Two PLP-dependent Enzymes

$A_{340}$ , and reaction rates were quantified using a molar extinction coefficient of  $6,220 \text{ M}^{-1}\cdot\text{cm}^{-1}$  for NADH. The substrate concentrations used for ALR-catalyzed assays ranged from  $(0.3\text{--}10) \times K_m$ , whereas those catalyzed by the CBL variants ranged from  $\leq 0.5$  to  $\geq 2 \times K_m$ . In the absence of enzyme, we found that the average rate of spontaneous NADH oxidation was  $< 0.5 \mu\text{M}\cdot\text{min}^{-1}$  (which was  $< 2\%$  of the maximal rate of ALR-catalyzed  $\text{NAD}^+$  formation). This background rate was measured and subtracted from each assay.

The continuous assay for cystathionine  $\beta$ -elimination was modified from Uren's protocol (62). Each reaction contained 50 mM Tris·HCl (pH 8.8) (or 100 mM Tris·HCl (pH 8.8), for assays with ALR variants), 0.4 mM DTNB, and enzyme. The final enzyme concentrations used in each reaction were as follows: CBL, 11 nM (in 1-ml assays); evolved CBL variants, 6.5–20.2 nM (in 1-ml assays); and ALR variants, 0.8–7  $\mu\text{M}$  (in 0.5-ml assays). Each reaction was pre-incubated at 37 °C for 2 min, and L-cystathionine was added to initiate the reaction. The thiol group of homocysteine (produced from the enzyme-catalyzed reaction) reacted with DTNB to form the 2-nitro-5-thiobenzoate dianion. The formation of 2-nitro-5-thiobenzoate dianion was monitored by an increase of  $A_{412}$  and was quantified using its molar extinction coefficient of  $14,150 \text{ M}^{-1}\cdot\text{cm}^{-1}$  (63). For CBL assays, the substrate concentrations typically ranged from  $(0.25\text{--}3) \times K_m$ , whereas those containing the evolved CBL variants ranged from 0.1 to  $\geq 1.6 \times K_m$ . Assays with the ALR(Y274F) variant contained substrate concentrations ranging from  $(0.6\text{--}3) \times K_m$ . The background rate of spontaneous DTNB decomposition was measured and subtracted from each assay.

Initial rate data were fitted to the Michaelis-Menten equation using Origin software (OriginLab; Northampton, MA). All Michaelis-Menten plots were constructed from activity data obtained at 7–9 substrate concentrations. The concentration of active enzyme in each assay was assumed to be the concentration of PLP-bound enzyme.

**Random Mutagenesis of *alr***—Random mutations were introduced into the *alr* gene using the GeneMorph II Random Mutagenesis kit from Agilent (La Jolla, CA). The epPCR had a total volume of 100  $\mu\text{l}$  and contained  $1 \times$  Mutazyme buffer, 2.5 units of Mutazyme DNA polymerase, 0.8 mM dNTPs, 0.5  $\mu\text{M}$  primer pCA24N.for, 0.5  $\mu\text{M}$  primer ALR.rev (5'-GCG GCC GCA TAG GCC TTA ATC CAC GTA TTT C-3'), and 570 ng of plasmid pCA24N-*alr* (corresponding to 100 ng of the *alr* template). The thermocycling conditions were 95 °C for 2 min, 35 cycles of 95 °C (30 s), 58 °C (30 s), 72 °C (70 s), and one final extension step at 72 °C for 2 min. The 1.2-kb amplified product was purified using a QIAquick spin column (Qiagen, Valencia, CA) and cloned into the SfiI sites of the pCA24N vector (23). *E. coli*  $\Delta\text{metC}$  cells were transformed with the ligated products by electroporation. Transformed cells were plated on LB-kanamycin-chloramphenicol library plates ( $245 \times 245$  mm, Corning) and allowed to grow at 37 °C for 15 h. The *alr* inserts from 20 randomly selected colonies were sequenced using primers pCA24N.for and pCA24N.rev2 (5'-CAA ATC CAG ATG GAG TTC TGA GG-3') to estimate the mutation spectrum of the library.

**Selection for ALR Variants with Higher CBL Activity**—Colonies harboring the randomly mutated *alr* variants were pooled by scraping them off the library plates in aliquots of LB medium, and then a 10- $\mu\text{l}$  aliquot of the pool was used to inoculate 50 ml of fresh LB/kanamycin/chloramphenicol medium. At  $A_{600} \sim 0.8$ , the cells were harvested by centrifugation at  $2,200 \times g$  and 4 °C for 15 min, followed by a wash with 50 ml of  $1 \times$  M9 medium. The cells were spread on an M9/glucose library plate supplemented with 5  $\mu\text{M}$  IPTG. For side-by-side comparison, *E. coli*  $\Delta\text{metC}$  cells carrying pCA24N-*alr* were spread on an additional plate. The same *E. coli* strain harboring an empty ASKA vector (pCA24N-NoIns (24)) was used as a negative control, whereas cells harboring pCA24N-*metC* were used as a positive control. After 2 days of incubation at 28 °C, a random subset of the colonies that had appeared on the library plate were picked. The randomly mutated *alr* allele from each colony was amplified by PCR with primers pCA24N.for and pCA24N.rev2, and the products were sequenced.

**Random Mutagenesis of *metC***—The *metC* gene was amplified in an epPCR, as described above for *alr* except that the primers used were pCA24N.for and *metC*(GFP-).rev (5'-GCG GCC GCA TAG GCC TTA TAC AAT TCG CGC A-3'), and the template was 1.3  $\mu\text{g}$  of plasmid pCA24N-*metC* (corresponding to 300 ng of the *metC* gene). The thermocycling conditions were 95 °C for 90 s, 30 cycles of 95 °C (30 s), 54 °C (30 s), 72 °C (90 s), and a final cycle of 72 °C for 60 s. The 1.2-kb product was digested with SfiI and ligated with the pCA24N(KanR) vector (a derivative of pCA24N, in which the chloramphenicol resistance marker had been replaced by a kanamycin marker), after it had been digested with the same restriction enzyme. The ligated products were used to transform *E. coli* DH5 $\alpha$ -E (Invitrogen) by electroporation, and transformed cells were grown on LB-kanamycin plates at 37 °C for 16 h. Fourteen colonies were picked at random for DNA sequencing to estimate the mutation spectrum of the library. The remainder of the transformed cells ( $\sim 1.5 \times 10^5$  colonies) were pooled, and the plasmids were isolated. The plasmid pool was used to transform *E. coli* MB2795. Transformed cells were spread on a library plate of LB-kanamycin supplemented with 0.5 mM D-Ala. Approximately  $3 \times 10^7$  colonies formed after 16 h of incubation at 37 °C. This ensured that 100% of the initial library was sampled (64).

**Selection for CBL Variants with Higher ALR Activity**—The *E. coli* MB2795 colonies carrying randomly mutated *metC* variants were pooled, and a 10- $\mu\text{l}$  aliquot was transferred into 50 ml of fresh LB-kanamycin that had been supplemented with D-Ala (0.5 mM). The cells were propagated at 37 °C to an  $A_{600}$  of 0.4. The culture was diluted appropriately, so that an estimated  $1.0 \times 10^6$  colony-forming units were spread on a library plate of LB-kanamycin supplemented with 5  $\mu\text{M}$  IPTG. After incubating the plate at 28 °C for 42 h, 48 random colonies were restreaked on fresh selective plates, both with and without 5  $\mu\text{M}$  IPTG, to confirm the growth phenotype of the clones. Clones with reproducible phenotypes were subjected to PCR screens and DNA sequencing analysis.

**Crystallization, Data Collection, and Structure Solution of ALR(Y274F) and CBL(P113S)**—Immediately prior to crystallization attempts with ALR(Y274F), a 20-ml spin concentrator

with nominal molecular mass cutoff of 50 kDa (GE Healthcare) was used to exchange the protein into a buffer consisting of 20 mM Tris-HCl (pH 8.0) and 100  $\mu$ M PLP. Protein was crystallized by vapor diffusion using the hanging drop method at a temperature of 291 K and with a reservoir solution (500  $\mu$ l) consisting of 1.6 M ammonium sulfate, 100 mM HEPES (pH 7.0). A 2:1 ratio of protein to crystallization buffer yielded yellow-colored hexagonal plate crystals that appeared overnight and grew to approximate dimensions of 0.4  $\times$  0.4  $\times$  0.1 mm over 7 days. After a week, crystals were supplemented with 10 mM PLP for 3 days to ensure complete cofactor incorporation.

Prior to data collection, an inhibited ALR(Y274F) structure was produced by soaking pre-formed crystals with 10 mM L-Ala-P. Inhibitor soaking was for 30 min, before a brief soak in crystallization buffer supplemented with 30% glycerol, and then flash-cooling in liquid nitrogen. Data were collected to 2.25 Å for the inhibitor-treated crystal and to 1.86 Å for an untreated crystal on a Rigaku MicroMax-007 HR rotating anode instrument equipped with a MAResearch Mar345dtb detector and Oxford Cryosystem Cobra cooling system. Data were processed using XDS (65) and scaled using SCALA (66). The structures were solved by molecular replacement using the wild type *E. coli* ALR structure (PDB 2RJG) and the program PHASER (67) and were modeled using COOT (68). The structures were refined using REFMAC5 (69) to *R*-factors of 17.5% (uninhibited) and 16.7% (inhibited), and free *R* values of 20.4 and 20.2%, respectively. The apparent space group of these crystals was *P*6<sub>22</sub>, but subsequently the true space group was determined as *P*6, with twin refinement in REFMAC5 improving the refinement statistics markedly. The final models were assessed by Ramachandran plot and geometrical analysis using both the MOLPROBITY server (70) and the PDB validation server. PLP and L-Ala-P-PLP aldimine molecules were modeled into discrete and unambiguous electron density, and sulfate and water molecules were added in the final rounds of refinement. The non-inhibited structure contains four protein molecules (two dimers), four PLPs, five sulfates, three glycerols, and 467 water molecules. The inhibited structure contains four protein molecules, three L-Ala-P-PLP aldimines, eight sulfates, and 236 water molecules. Active site electron density is shown in the supplemental Fig. S1, and data collection and refinement statistics are in supplemental Table S2.

The CBL(P113S) protein was treated similarly to ALR (Y274F), apart from the following specific details. Buffer exchange prior to crystallization used 5 mM HEPES (pH 7.5) supplemented with 10  $\mu$ M PLP. Protein was finally concentrated to 7.5 mg·ml<sup>-1</sup> and crystallized by vapor diffusion using the hanging drop method and with a reservoir solution consisting of 10% PEG400, 100 mM HEPES (pH 7.3), and 150 mM CaCl<sub>2</sub>. Various ratios of protein to reservoir solution were mixed (1 + 1  $\mu$ l, 2 + 2  $\mu$ l, 3 + 1.5  $\mu$ l, and 1.5 + 3  $\mu$ l) and inverted over the reservoir solution. In all conditions, yellow-colored coffin-shaped crystals grew overnight, with the largest reaching a final size of  $\sim$ 0.2  $\times$  0.2  $\times$  0.4 mm within 2 days. A large and perfect looking crystal was coated in a 70:30 mixture of paratone N and mineral oil before flash-cooling in liquid nitrogen; it was diffracted to a resolution of 1.74 Å. The structure was solved by molecular replacement using the wild type

CBL structure (PDB 1CL1). The P113S mutation was incorporated, and a HEPES molecule plus 379 water molecules were included in the final model. The PLP cofactor was clearly observed in the structure at full occupancy.

An inhibited CBL(P113S) structure was produced by soaking a pre-formed crystal overnight in a crystallization solution supplemented with 10 mM L-Ala-P. Data were collected in-house and processed, and the structure was solved and refined as outlined above. The L-Ala-P-PLP aldimine was clearly observed in discrete and unambiguous electron density and was included in final refinement along with 514 water molecules. Active site electron density is shown in supplemental Fig. S1, and data collection and refinement statistics for the uninhibited and inhibited CBL(P113S) structures are in supplemental Table S2.

**Modeling a Cystathionine-PLP Adduct into the ALR(Y274F) Active Site**—A cystathionine-PLP adduct was built using the ligand building module in COOT (68). This molecule was docked into our crystal structure model of ALR(Y274F); to prepare the protein model, all ligands and water molecules were removed and hydrogen atoms added. Docking used the program GOLD version 5.2 (CCDC); the GOLD wizard and default docking parameters were used to produce an initial command file that was subsequently edited to include a scaffold match constraint. The PLP fragment of the L-Ala-P-PLP molecule in our inhibited ALR(Y274F) crystal structure was used as the scaffold, essentially forcing the PLP of the cystathionine-PLP adduct to assume the inhibited crystal structure position. The cystathionine fragment was allowed to rotate freely. Atom typing was assigned for both the protein and cystathionine-PLP molecules by GOLD, and 10 docking solutions were produced for manual inspection.

**Author Contributions**—W. M. P. conceived and coordinated the study. V. W. C. S. designed and performed all of the molecular biology, enzyme kinetics and directed evolution experiments. Y. Y. and C. J. S. designed and conducted the crystallographic experiments. C. J. S. carried out the docking analysis. V. W. C. S., C. J. S., and W. M. P. analyzed the data and wrote the paper. All authors reviewed the results and approved the final version of the manuscript.

**Acknowledgments**—We thank Dr. Monica Gerth (University of Otago) for the *pCA24N(KanR)* vector. We are also grateful to Natasha Comer (Massey University) and James McKellar (University of Otago) for technical assistance. *E. coli* strain MB2795 was a kind gift from Prof. Kurt Krause (University of Otago).

## References

1. Percudani, R., and Peracchi, A. (2009) The B6 database: a tool for the description and classification of vitamin B6-dependent enzymatic activities and of the corresponding protein families. *BMC Bioinformatics* **10**, 273
2. Raboni, S., Spyarakis, F., Campanini, B., Amadasi, A., Bettati, S., Peracchi, A., Mozzarelli, A., and Contestabile, R. (2010) in *Comprehensive Natural Products II: Chemistry and Biology* (Mander, L. and Liu, H.-W., editors in chief) pp. 273–350, Elsevier Ltd., New York
3. Percudani, R., and Peracchi, A. (2003) A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Rep.* **4**, 850–854
4. Toney, M. D. (2011) Controlling reaction specificity in pyridoxal phosphate enzymes. *Biochim. Biophys. Acta* **1814**, 1407–1418
5. Richard, J. P., Amyes, T. L., Crugeiras, J., and Rios, A. (2011) The PLP cofactor: lessons from studies on model reactions. *Biochim. Biophys. Acta* **1814**, 1419–1425

## Reciprocal Promiscuity in Two PLP-dependent Enzymes

- Grishin, N. V., Phillips, M. A., and Goldsmith, E. J. (1995) Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Sci.* **4**, 1291–1304
- Schneider, G., Käck, H., and Lindqvist, Y. (2000) The manifold of vitamin B6 dependent enzymes. *Structure* **8**, R1–R6
- Denesyuk, A. I., Denessiouk, K. A., Korpela, T., and Johnson, M. S. (2002) Functional attributes of the phosphate group binding cup of pyridoxal phosphate-dependent enzymes. *J. Mol. Biol.* **316**, 155–172
- Setiyaputra, S., Mackay, J. P., and Patrick, W. M. (2011) The structure of a truncated phosphoribosylanthranilate isomerase suggests a unified model for evolution of the ( $\beta\alpha$ )<sub>8</sub> barrel fold. *J. Mol. Biol.* **408**, 291–303
- Patrick, W. M., Quandt, E. M., Swartzlander, D. B., and Matsumura, I. (2007) Multicopy suppression underpins metabolic evolvability. *Mol. Biol. Evol.* **24**, 2716–2722
- Clausen, T., Huber, R., Laber, B., Pohlenz, H.-D., and Messerschmidt, A. (1996) Crystal structure of the pyridoxal-5'-phosphate dependent cystathionine  $\beta$ -lyase from *Escherichia coli* at 1.83 Å. *J. Mol. Biol.* **262**, 202–224
- Ejim, L. J., Blanchard, J. E., Koteva, K. P., Sumerfield, R., Elowe, N. H., Chechetto, J. D., Brown, E. D., Junop, M. S., and Wright, G. D. (2007) Inhibitors of bacterial cystathionine  $\beta$ -lyase: leads for new antimicrobial agents and probes of enzyme structure and function. *J. Med. Chem.* **50**, 755–764
- Anthony, K. G., Strych, U., Yeung, K. R., Shoen, C. S., Perez, O., Krause, K. L., Cynamon, M. H., Aristoff, P. A., and Koski, R. A. (2011) New classes of alanine racemase inhibitors identified by high-throughput screening show antimicrobial activity against *Mycobacterium tuberculosis*. *PLoS ONE* **6**, e20374
- Ciustea, M., Mootien, S., Rosato, A. E., Perez, O., Cirillo, P., Yeung, K. R., Ledizet, M., Cynamon, M. H., Aristoff, P. A., Koski, R. A., Kaplan, P. A., and Anthony, K. G. (2012) Thiadiazolidinones: a new class of alanine racemase inhibitors with antimicrobial activity against methicillin-resistant *Staphylococcus aureus*. *Biochem. Pharmacol.* **83**, 368–377
- Wu, D., Hu, T., Zhang, L., Chen, J., Du, J., Ding, J., Jiang, H., and Shen, X. (2008) Residues Asp164 and Glu165 at the substrate entryway function potentially in substrate orientation of alanine racemase from *E. coli*: enzymatic characterization with crystal structure analysis. *Protein Sci.* **17**, 1066–1076
- Shaw, J. P., Petsko, G. A., and Ringe, D. (1997) Determination of the structure of alanine racemase from *Bacillus stearothermophilus* at 1.9-Å resolution. *Biochemistry* **36**, 1329–1342
- Stamper, G. F., Morollo, A. A., Ringe, D., and Stamper, C. G. (1998) Reaction of alanine racemase with 1-aminoethylphosphonic acid forms a stable external aldimine. *Biochemistry* **37**, 10438–10445
- Watanabe, A., Kurokawa, Y., Yoshimura, T., Kurihara, T., Soda, K., Esaki, N., and Watababe, A. (1999) Role of lysine 39 of alanine racemase from *Bacillus stearothermophilus* that binds pyridoxal 5'-phosphate. Chemical rescue studies of Lys-39  $\rightarrow$  Ala mutant. *J. Biol. Chem.* **274**, 4189–4194
- Watanabe, A., Yoshimura, T., Mikami, B., and Esaki, N. (1999) Tyrosine 265 of alanine racemase serves as a base abstracting  $\alpha$ -hydrogen from L-alanine: the counterpart residue to lysine 39 specific to D-alanine. *J. Biochem.* **126**, 781–786
- Ondrechen, M. J., Briggs, J. M., and McCammon, J. A. (2001) A model for enzyme-substrate interaction in alanine racemase. *J. Am. Chem. Soc.* **123**, 2830–2834
- Watanabe, A., Yoshimura, T., Mikami, B., Hayashi, H., Kagamiyama, H., and Esaki, N. (2002) Reaction mechanism of alanine racemase from *Bacillus stearothermophilus*: x-ray crystallographic studies of the enzyme bound with N-(5'-phosphopyridoxyl)alanine. *J. Biol. Chem.* **277**, 19166–19172
- Spies, M. A., and Toney, M. D. (2003) Multiple hydrogen kinetic isotope effects for enzymes catalyzing exchange with solvent: application to alanine racemase. *Biochemistry* **42**, 5099–5107
- Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005) Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.* **12**, 291–299
- Soo, V. W., Hanson-Manful, P., and Patrick, W. M. (2011) Artificial gene amplification reveals an abundance of promiscuous resistance determinants in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1484–1489
- Kim, J., Kershner, J. P., Novikov, Y., Shoemaker, R. K., and Copley, S. D. (2010) Three serendipitous pathways in *E. coli* can bypass a block in pyridoxal-5'-phosphate synthesis. *Mol. Syst. Biol.* **6**, 436
- Patrick, W. M., and Matsumura, I. (2008) A study in molecular contingency: glutamine phosphoribosylpyrophosphate amidotransferase is a promiscuous and evolvable phosphoribosylanthranilate isomerase. *J. Mol. Biol.* **377**, 323–336
- Peisajovich, S. G., and Tawfik, D. S. (2007) Protein engineers turned evolutionists. *Nat. Methods* **4**, 991–994
- Firth, A. E., and Patrick, W. M. (2008) GLUE-IT and PEDEL-AA: new programmes for analyzing protein diversity in randomized libraries. *Nucleic Acids Res.* **36**, W281–W285
- LeMagueres, P., Im, H., Ebalunode, J., Strych, U., Benedik, M. J., Briggs, J. M., Kohn, H., and Krause, K. L. (2005) The 1.9 Å crystal structure of alanine racemase from *Mycobacterium tuberculosis* contains a conserved entryway into the active site. *Biochemistry* **44**, 1471–1481
- Scaletti, E. R., Luckner, S. R., and Krause, K. L. (2012) Structural features and kinetic characterization of alanine racemase from *Staphylococcus aureus* (Mu50). *Acta Crystallogr. D Biol. Crystallogr.* **68**, 82–92
- Somarowthu, S., and Ondrechen, M. J. (2012) POOL server: machine learning application for functional site prediction in proteins. *Bioinformatics* **28**, 2078–2079
- Yow, G.-Y., Watanabe, A., Yoshimura, T., and Esaki, N. (2003) Conversion of the catalytic specificity of alanine racemase to a D-amino acid aminotransferase activity by a double active-site mutation. *J. Mol. Catal. B Enzym.* **23**, 311–319
- Seebeck, F. P., and Hilvert, D. (2003) Conversion of a PLP-dependent racemase into an aldolase by a single active site mutation. *J. Am. Chem. Soc.* **125**, 10158–10159
- Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2256–2268
- Clausen, T., Huber, R., Messerschmidt, A., Pohlenz, H. D., and Laber, B. (1997) Slow-binding inhibition of *Escherichia coli* cystathionine  $\beta$ -lyase by L-aminoethoxyvinylglycine: a kinetic and x-ray study. *Biochemistry* **36**, 12633–12643
- Clausen, T., Laber, B., and Messerschmidt, A. (1997) Mode of action of cystathionine  $\beta$ -lyase. *Biol. Chem.* **378**, 321–326
- Strych, U., Penland, R. L., Jimenez, M., Krause, K. L., and Benedik, M. J. (2001) Characterization of the alanine racemases from two mycobacteria. *FEMS Microbiol. Lett.* **196**, 93–98
- Kang, L., Shaw, A. C., Xu, D., Xia, W., Zhang, J., Deng, J., Wöldike, H. F., Liu, Y., and Su, J. (2011) Upregulation of MetC is essential for D-alanine-independent growth of an *abr/dadX*-deficient *Escherichia coli* strain. *J. Bacteriol.* **193**, 1098–1106
- Lodha, P. H., and Aitken, S. M. (2011) Characterization of the side-chain hydroxyl moieties of residues Y56, Y111, Y238, Y338, and S339 as determinants of specificity in *E. coli* cystathionine  $\beta$ -lyase. *Biochemistry* **50**, 9876–9885
- Søndergaard, C. R., Olsson, M. H., Rostkowski, M., and Jensen, J. H. (2011) Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pK<sub>a</sub> values. *J. Chem. Theory Comput.* **7**, 2284–2295
- Jensen, R. A. (1976) Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* **30**, 409–425
- Glasner, M. E., Gerlt, J. A., and Babbitt, P. C. (2006) Evolution of enzyme superfamilies. *Curr. Opin. Chem. Biol.* **10**, 492–497
- Khersonsky, O., and Tawfik, D. S. (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **79**, 471–505
- Copley, S. D. (2015) An evolutionary biochemist's perspective on promiscuity. *Trends Biochem. Sci.* **40**, 72–78
- Spies, M. A., Woodward, J. J., Watnik, M. R., and Toney, M. D. (2004) Alanine racemase free energy profiles from global analyses of progress curves. *J. Am. Chem. Soc.* **126**, 7464–7475
- Patrick, W. M., Weisner, J., and Blackburn, J. M. (2002) Site-directed mutagenesis of Tyr354 in *Geobacillus stearothermophilus* alanine racemase identifies a role in controlling substrate specificity and a possible role in the evolution of antibiotic resistance. *ChemBioChem* **3**, 789–792

47. Lodha, P. H., Jaworski, A. F., and Aitken, S. M. (2010) Characterization of site-directed mutants of residues R58, R59, D116, W340 and R372 in the active site of *E. coli* cystathionine  $\beta$ -lyase. *Protein Sci.* **19**, 383–391
48. McCune, C. D., Chan, S. J., Beio, M. L., Shen, W., Chung, W. J., Szczesniak, L. M., Chai, C., Koh, S. Q., Wong, P. T., and Berkowitz, D. B. (2016) “Zipped synthesis” by cross-metathesis provides a cystathionine  $\beta$ -synthase inhibitor that attenuates cellular H<sub>2</sub>S levels and reduces neuronal infarction in a rat ischemic stroke model. *ACS Cent. Sci.* **2**, 242–252
49. Mehta, P. K., and Christen, P. (2000) The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *Adv. Enzymol. Relat. Areas Mol. Biol.* **74**, 129–184
50. Eliot, A. C., and Kirsch, J. F. (2004) Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem.* **73**, 383–415
51. Newton, M. S., Arcus, V. L., and Patrick, W. M. (2015) Rapid bursts and slow declines: on the possible evolutionary trajectories of enzymes. *J. R. Soc. Interface* **12**, 20150036
52. di Salvo, M. L., Florio, R., Paiardini, A., Vivoli, M., D'Aguanno, S., and Contestabile, R. (2013) Alanine racemase from *Tolypocladium inflatum*: a key PLP-dependent enzyme in cyclosporin biosynthesis and a model of catalytic promiscuity. *Arch. Biochem. Biophys.* **529**, 55–65
53. Wolf, Y. I., and Koonin, E. V. (2013) Genome reduction as the dominant mode of evolution. *BioEssays* **35**, 829–837
54. Adams, N. E., Thiaville, J. J., Proestos, J., Juárez-Vázquez, A. L., McCoy, A. J., Barona-Gómez, F., Iwata-Reuyl, D., de Crécy-Lagard, V., and Maurilli, A. T. (2014) Promiscuous and adaptable enzymes fill “holes” in the tetrahydrofolate pathway in *Chlamydia* species. *mBio* **5**, e01378–01314
55. Bertoldi, M., Cellini, B., Paiardini, A., Di Salvo, M., and Borri Voltattorni, C. (2003) *Treponema denticola* cystathionine exhibits significant alanine racemase activity accompanied by transamination: mechanistic implications. *Biochem. J.* **371**, 473–483
56. Vollmer, J., Schiefer, A., Schneider, T., Jülicher, K., Johnston, K. L., Taylor, M. J., Sahl, H. G., Hoerauf, A., and Pfarr, K. (2013) Requirement of lipid II biosynthesis for cell division in cell wall-less *Wolbachia*, endobacteria of arthropods and filarial nematodes. *Int. J. Med. Microbiol.* **303**, 140–149
57. Foster, J., Ganatra, M., Kamal, I., Ware, J., Makarova, K., Ivanova, N., Bhattacharyya, A., Kapatral, V., Kumar, S., Posfai, J., Vincze, T., Ingram, J., Moran, L., Lapidus, A., Omelchenko, M., et al. (2005) The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. *PLoS Biol.* **3**, e121
58. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008
59. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423
60. Inagaki, K., Tanizawa, K., Badet, B., Walsh, C. T., Tanaka, H., and Soda, K. (1986) Thermostable alanine racemase from *Bacillus stearothermophilus*: molecular cloning of the gene, enzyme purification, and characterization. *Biochemistry* **25**, 3268–3274
61. Esaki, N., and Walsh, C. T. (1986) Biosynthetic alanine racemase of *Salmonella typhimurium*: purification and characterization of the enzyme encoded by the *alr* gene. *Biochemistry* **25**, 3261–3267
62. Uren, J. R. (1987) Cystathionine  $\beta$ -lyase from *Escherichia coli*. *Methods Enzymol.* **143**, 483–486
63. Riddles, P. W., Blakeley, R. L., and Zerner, B. (1979) Ellman's reagent: 5,5'-dithiobis(2-nitrobenzoic acid)—a reexamination. *Anal. Biochem.* **94**, 75–81
64. Firth, A. E., and Patrick, W. M. (2005) Statistics of protein library construction. *Bioinformatics* **21**, 3314–3315
65. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132
66. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82
67. 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
68. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
69. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367
70. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21