Chemistry & Biology 13, 1317–1326, December 2006 DOI 10.1016/j.chembiol.2006.10.008

Discovery of a Substrate Selectivity Switch in Tyrosine Ammonia-Lyase, a Member of the Aromatic Amino Acid Lyase Family

Kevin T. Watts,[1] Benjamin N. Mijts,[1] Pyung Cheon Lee,[1] Andrew J. Manning,[1] and Claudia Schmidt-Dannert[1,2,*]

[1] Department of Biochemistry, Molecular Biology, and Biophysics
University of Minnesota
1479 Gortner Avenue
St. Paul, Minnesota 55108

Summary

Tyrosine ammonia-lyase (TAL) is a recently described member of the aromatic amino acid lyase family, which also includes phenylalanine (PAL) and histidine ammonia-lyases (HAL). TAL is highly selective for L-tyrosine, and synthesizes 4-coumaric acid as a protein cofactor or antibiotic precursor in microorganisms. In this report, we identify a single active site residue important for substrate selection in this enzyme family. Replacing the active site residue His89 with Phe in TAL completely switched its substrate selectivity from tyrosine to phenylalanine, thereby converting it into a highly active PAL. When a corresponding mutation was made in PAL, the enzyme lost PAL activity and gained TAL activity. The discovered substrate selectivity switch is a rare example of a complete alteration of substrate specificity by a single point mutation. We also show that the identity of the amino acid at the switch position can serve as a guide to predict substrate specificities of annotated aromatic amino acid lyases in genome sequences.

Introduction

Tyrosine ammonia-lyase (TAL, EC 4.3.1.5) has recently been described as a member of the aromatic amino acid lyase family of enzymes, which also includes phenylalanine ammonia-lyase (PAL, EC 4.3.1.5) and histidine ammonia-lyase (HAL, EC 4.3.1.3). These enzymes catalyze the nonoxidative deamination of their cognate aromatic L-amino acid substrate to a *trans*-2,3-unsaturated substituted propenoic acid and ammonia [1] (Figure 1A). HALs and PALs are fairly ubiquitous in nature, but are involved in several different processes. HAL catalyzes the first step in the degradation of histidine, deaminating histidine to produce urocanic acid, which is further metabolized to glutamate [1]. PALs from bacteria and plants deaminate phenylalanine to cinnamic acid, which is ultimately converted into secondary metabolites such as coumarins and flavonoids in plants [2] and several antibiotic compounds in bacteria [3, 4]. In humans, recombinant PAL is currently being explored for treatment of phenylketonuria (PKU) by metabolizing excess dietary phenylalanine [5]. In contrast to HAL and PAL, TAL is quite rare, identified thus far only in *Rhodobacter capsulatus* (Integrated Genomics accession number: RRC01844) [6], *Rba. sphaeroides* (GenBank accession number: YP_355075) [7], and, very recently, *Saccharothrix espanaensis* (GenBank accession number: ABC88669) [8]. This rarity is likely a reflection of its specialized role in 4-coumaric acid biosynthesis, which is used as the cofactor for photoactive yellow protein (PYP) in *Rhodobacter* [9], and initiates the conversion of tyrosine to the *N*-(*m*,*p*-dihydroxycinnamoyl)-taurine moiety of saccharomicin antibiotics [8].

Unique to the aromatic amino acid lyases is the modified amino acid cofactor 3,5-dihydro-5-methylidene-4*H*-imidazol-4-one (MIO), formed autocatalytically during nascent peptide folding from a highly conserved Ala-Ser-Gly motif [10] even when expressed in recombinant *Escherichia coli* [11]. This highly electrophilic cofactor, first identified in HAL [12], is not found in other members of the L-amino acid lyase family, such as aspartate ammonia-lyase [13]. The exact mechanism of ammonia elimination in the aromatic amino acid ammonia-lyases is currently unknown, although two mechanisms have been proposed based upon experimental data [1]. The recently solved crystal structures of PAL from two different sources (*Rhodosporidium toruloides* PAL [14] and *Petroselinum crispum* PAL1 [15]) illustrate the conflicting mechanistic proposals. In one mechanism, electrophilic attack of MIO on the substrate aromatic ring results in loss of aromaticity, and generates a delocalized carbocation within the benzyl ring system of enzyme-bound substrate. The direct result of this Friedel-Crafts-like acylation reaction is increased acidity for the substrate β proton, whose removal by an enzyme general base facilitates the electronic rearrangement that ultimately results in elimination of ammonia and double-bond formation between the α and β carbons [1, 16]. The other mechanism, called the E_1cB mechanism, involves nucleophilic attack on the substrate α-amino group by MIO [17, 18]. An enzymatic general base then removes the β proton, creating a carbanion intermediate, which collapses, resulting in elimination of the α-amino group and formation of the double bond between α and β carbons. Interpretation of structural data for PAL from *R. toruloides* suggested the latter mechanistic route [14]. It was even suggested that a His found within the active site of PAL (His137) was responsible for the general base catalysis to remove the β proton. This His is also present in the amino acid sequence of TAL (His89). However, not all PALs have this supposedly essential catalytic residue, leading to speculation about its functional significance.

Each of the aromatic ammonia-lyases is quite specific for their individual substrates. TAL from *Rba. capsulatus*, for example, has a 150-fold higher catalytic efficiency (k_{cat}/K_m) for tyrosine over the similar substrate phenylalanine [6], while parsley PAL1 has been demonstrated to have a 10,000-fold higher catalytic efficiency for phenylalanine than tyrosine [19], and similar results were observed with *Arabidopsis* PAL isoforms [20]. However, there appears to be another class of bifunctional PALs that has the ability to use tyrosine and phenylalanine with almost equal efficiency. This phenomenon was

*Correspondence: schmi232@umn.edu
[2] Lab address: http://www.cbs.umn.edu/bmbb/faculty/csd/

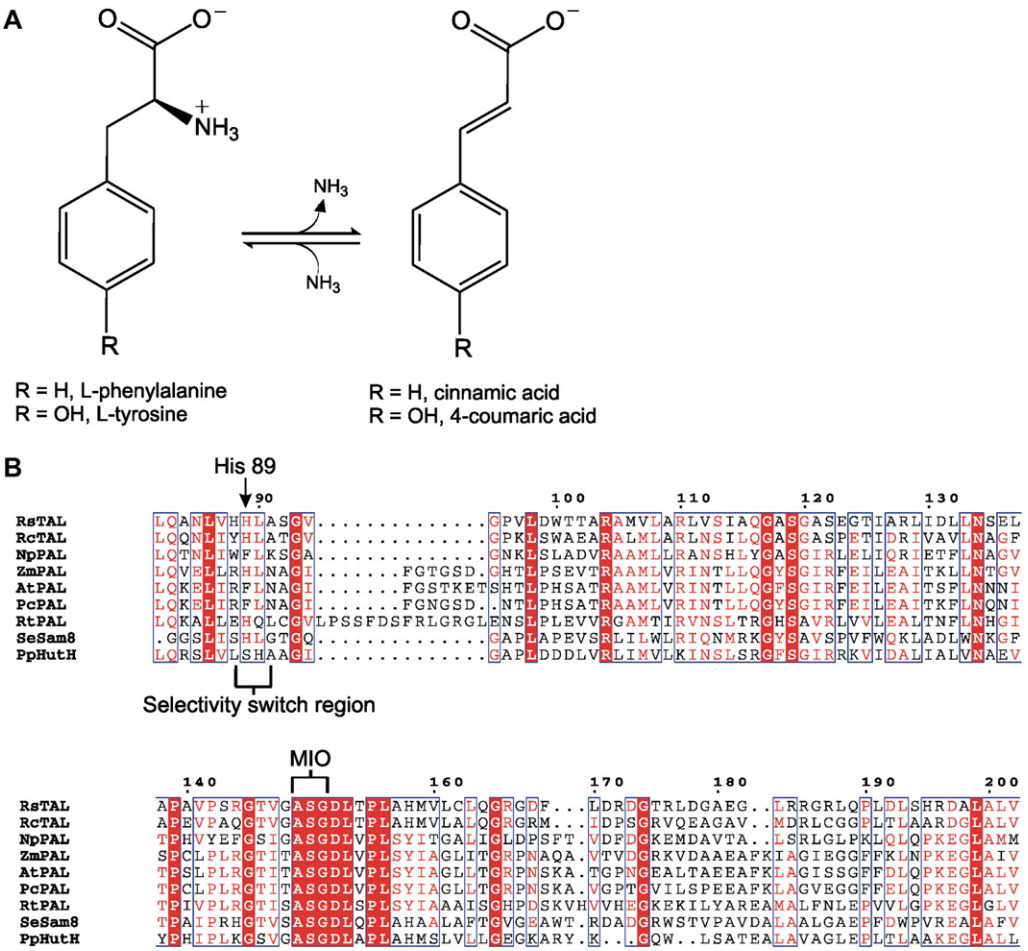

Figure 1. Reaction and Sequence Alignment of Phenylalanine and Tyrosine Ammonia-Lyases

(A) Reaction of PALs and TALs. The α-amino group is removed from the aromatic L-amino acid substrate to create an α,β-unsaturated phenylpropenoic acid and ammonia. Deamination of tyrosine produces 4-coumaric acid, while phenylalanine deamination results in formation of cinnamic acid.

(B) Sequence alignment of PALs and TALs. Shown is the location of the active site His residue (His89) critical in substrate selection and the three residues (A149-S150-G151) that comprise the posttranslationally modified MIO cofactor. Shaded residues are conserved across all aligned sequences, and boxed residues are conserved across at least four sequences or share physical properties without sequence identity. Sequences were aligned using the Multalin sequence alignment tool. Abbreviations and accession numbers are: RsTAL, *Rba. sphaeroides* TAL (GenBank accession number: YP_355075); RcTAL, *Rba. capsulatus* TAL (Integrated Genomics accession number: RRC01844); NpPAL, *N. punctiforme* PCC 73102 PAL (GenBank accession number: ZP_00105927); ZmPAL, *Z. mays* PAL (GenBank accession number: AAL40137); PcPAL, *P. crispum* PAL1 (GenBank accession number: CAA68938); AtPAL, *A. thaliana* PAL1 (GenBank accession number: AAP59438); RtPAL, *R. toruloides* PAL (GenBank accession number: CAA35886); SeSam8, *S. espanaensis* Sam8 (GenBank accession number: ABC88669); PpHutH, *P. putida* HutH (GenBank accession number: P21310).

most clearly observed in the PAL characterized from *Zea mays* [21]. These bifunctional PALs seem to be common among monocotyledonous plants, such as corn, wheat, and rice [22], as well as yeasts [23, 24], and in particular the yeast *R. toruloides* (formerly *Rhodotorula glutinis*) [18, 25]. The mechanism of how these bifunctional PALs are able to use both substrates, while other PALs and TALs are highly specific for their respective substrates, is currently not understood.

We previously identified a TAL from *Rba. sphaeroides* [7] that, when expressed in *E. coli*, produced only 4-coumaric acid from tyrosine in vivo. In this report, we seek to further characterize the activity of this TAL from *Rba. sphaeroides* in vitro, and determine how this enzyme is able to discriminate so effectively between the two structurally similar substrates tyrosine and phenylalanine. Through comparative sequence analysis and protein modeling, we identified a single amino acid position in the TAL active site that controls substrate selection. When mutated, this single amino acid functions as a highly efficient substrate selectivity switch, converting the enzyme from exclusively TAL activity to exclusively PAL activity, which makes this a rare example of a single amino acid almost completely dictating the substrate preference of a particular enzyme. The corresponding mutation made in PAL1 from *Arabidopsis thaliana* resulted in lowered PAL activity and increased TAL activity. Finally, using the identified single amino acid position as a predictor of activity, a PAL from *Nostoc punctiforme* PCC 73102 was identified and cloned, which makes it, to our knowledge, the first PAL described from a cyanobacterium.

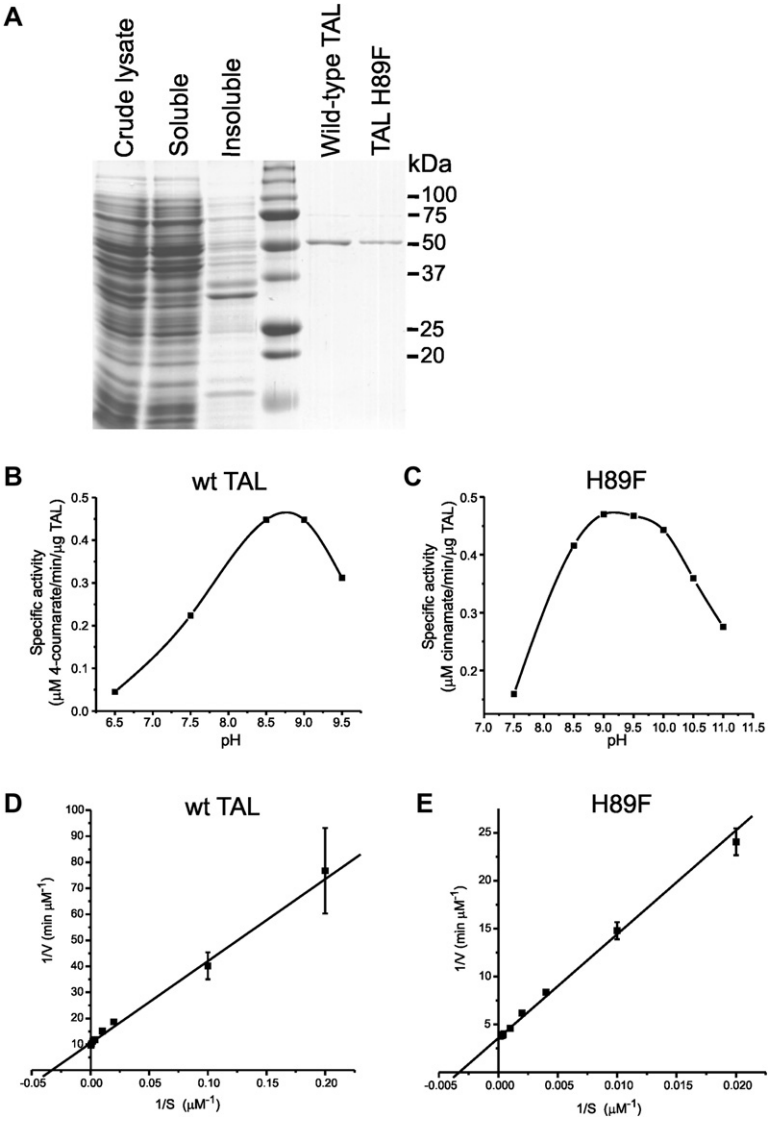

Figure 2. Comparison of Wild-Type TAL and TAL H89F

(A) SDS-PAGE gel (12%) showing purification of wild-type TAL and TAL H89F from *E. coli*. Lane 1 is crude cell lysate, lane 2 is soluble protein fraction, and lane 3 is insoluble protein fraction from TAL_NHis culture. Purification of TAL H89F was similar. Lane 4 is protein standards with corresponding molecular weights shown at right. Lane 5 is purified wild-type TAL and lane 6 is purified TAL H89F.

(B) pH optimum of wild-type TAL. Optimal TAL activity was determined with tyrosine over the pH range of 6.5–9.5. Optimal activity was determined to be between 8.5 and 9.0.

(C) pH optimum of TAL H89F. Optimal TAL H89F activity was determined with phenylalanine over the pH range of 7.5–11.0. Optimal activity with this substrate was observed at pH 9.0.

(D) Double reciprocal plot of wild-type TAL with tyrosine. Substrate concentration was varied over the range of 5–2000 μM and 25 μg TAL was used.

(E) Double reciprocal plot of TAL H89F with L-phenylalanine. Substrate concentration was varied over the range of 50–5000 μM and 10 μg TAL H89F was used. Data points represent the mean of six independent measurements ± standard deviation.

Results

Sequence Comparison and Catalytic Properties of TAL

TAL from *Rba. sphaeroides* was previously expressed in *E. coli* and demonstrated high TAL activity in vivo, as 4-coumaric acid production was observed. TAL was then used to engineer a hybrid flavonoid pathway in *E. coli* [7]. This TAL shows the highest homology (56% amino acid identity) to a previously reported TAL from *Rba. capsulatus* [6]. However, TAL exhibits only moderate to low protein sequence identity (ranging between 25% and 35%) with HALs and PALs. The enzyme's identity to the two crystallized PALs from *P. crispum* (PDB code: 1W27) and *R. toruloides* (PDB code: 1T6P) is 29% and 31%, respectively, while TAL shows a slightly higher identity of 37% with the crystallized HAL from *P. putida* (PDB code: 1B8F). Key active site residues and, most importantly, the sequence motif Ala-Ser-Gly for the MIO cofactor [12, 14, 15], are conserved in TAL (Figure 1B).

Because we previously established TAL's selectivity for tyrosine only in vivo in *E. coli*, the enzyme was N-terminally His tagged for affinity purification and its activity was characterized. Purified TAL ran at the molecular mass of 55 kDa and was 90% pure as estimated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 2A). The pH optimum was found to be near 8.5–9.0 (Figure 2B), which agrees well with pH optima reported for *Rba. capsulatus* TAL and other PALs from microbial and plant sources [6, 20, 26]. A kinetic study of TAL was then carried out with both tyrosine and phenylalanine as substrates. As seen in Table 1, TAL showed a strong preference for tyrosine compared to phenylalanine. The k_{cat} observed for both substrates, however, was very similar, indicating that both substrates can be turned over almost equally well. The effect of divalent metals on TAL activity was also investigated. Contrary to previous data showing metal activation in HAL [27], no rate enhancement was seen in the presence of divalent metals, nor was there significant enzyme inhibition in the presence of increasing ethylenediaminetetraacetic acid (EDTA) concentration (data not shown).

Table 1. Steady-State Kinetic Parameters of TALs and PALs Investigated in This Study (Underlined) with Tyrosine and Phenylalanine as Substrates

Organism	Residue at Switch Position	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
Tyrosine				
Rba. sphaeroides TAL[a]	H	31.4 ± 8.8	3.4 ± 0.15	1.1×10^5
Rba. capsulatus TAL[b]	H	15.6	27.7	1.8×10^6
Z. mays PAL[c]	H	19 ± 1	0.92 ± 0.22	4.8×10^4
A. thaliana PAL1 F144H[a]	H	532.6 ± 27.7	0.75 ± 0.15	1410
Rba. sphaeroides TAL H89F[a]	F	1541 ± 845.1	$0.0087 \pm 9.2 \times 10^{-4}$	5.6
P. crispum PAL1[d]	F	2500 ± 300	0.3	120
A. thaliana PAL1[a]	F	1049 ± 67.2	0.084 ± 0.0038	80
Phenylalanine				
Rba. sphaeroides TAL[a]	H	5204 ± 1441	2.1 ± 0.45	403
Rba. capsulatus TAL[b]	H	1277	15.1	1.2×10^4
Z. mays PAL[c]	H	270 ± 20	10.6 ± 3.2	3.9×10^4
A. thaliana PAL1 F144H[a]	H	1386 ± 162.9	47.3 ± 2.4	3.4×10^4
Rba. sphaeroides TAL H89F[a]	F	269.1 ± 42.0	24.0 ± 0.089	8.9×10^4
P. crispum PAL1[d]	F	17.2 ± 0.7	22.0	1.3×10^6
A. thaliana PAL1[a]	F	42.8 ± 1.3	115.8 ± 2.1	2.7×10^6

Shown for comparison are kinetic data from enzymes reported in the literature. Data for underlined enzymes represent the mean of six independent experiments ± standard deviation.
[a] From this study.
[b] From [6].
[c] From [21].
[d] From [19].

Identification of His89 as Putative Substrate Selectivity Switch in TAL

After confirming TAL's selectivity for tyrosine as substrate, we set out to identify the amino acid residue(s) responsible for efficient discrimination between tyrosine and phenylalanine. A sequence alignment was created (Figure 1B) with the three known TALs, several PALs, and the well-characterized HAL from *P. putida* (HutH, GenBank accession number: P21310) [12]. The PALs were chosen based on available structural information [14, 15] and their established ability to use either phenylalanine exclusively or both phenylalanine and tyrosine as substrates [19–21, 25]. A model structure of TAL was also created based upon the structure of *R. toruloides* PAL, which displayed good conservation of the active site and overall structure, with a root-mean-square deviation (rmsd) of 2.9 Å over the most similar 251 amino acids (Figure 3).

Examination of the sequence alignment revealed a single amino acid (His89, TAL numbering) that was not absolutely conserved within the active site defined in the structures of HAL and PAL and the model of TAL (Figures 1B and 3). This residue varied between a Phe or His in PAL sequences used for alignment, while all three TAL sequences, and those PALs that also demonstrate increased TAL activity, have His at the identified position. Those PALs that display a strong specificity for phenylalanine have a Phe at the same position. In HAL, the equivalent His residue (His83) is found within interacting distance of the imidazole ring when the substrate is modeled using Friedel-Crafts conditions [12]. This residue has also been proposed to be involved in binding and orientation of the substrate in HAL [28], but how the corresponding residue interacts with the substrates L-tyrosine and L-phenylalanine within the context of the TAL active site is not known. Previous work has suggested that amino acids in this region may be important in substrate selection. An H83L mutation in HAL resulted in inactive enzyme [28], and an attempt to convert *P. crispum* PAL to HAL by mutation of the adjacent Leu (Leu138) residue resulted in PAL with a 100-fold increase in K_m for phenylalanine and no difference in K_m for histidine [29].

The His/Phe amino acid position identified through sequence alignment was examined in the available PAL structures, and compared to the model structure of TAL. The Phe residue located in the *P. crispum* PAL (PDB code: 1W27) (Phe137) points into the active site cavity and is approximately 10 Å from the methylidene carbon of MIO, with an orientation that is slightly off perpendicular to the vertical plane of MIO. In the structure of *R. toruloides* PAL (PDB code: 1T6P), the His occupies an almost isosteric position, and is approximately 8.5 Å away from the methylidene carbon of MIO. The TAL model has the His in approximately the same position, but is 9.2 Å from the MIO methylidene carbon (Figure 3). In both PAL structures and the TAL model, this residue lies in the same horizontal plane as a conserved Arg residue, believed to be involved in stabilization of the negatively charged carboxylate of the substrate [12]. The distance from the distal Arg to the Phe or His residue is 14–15 Å, which appears to define an active site cavity with Arg at the carboxylate end, and the Phe or His residue at the end nearest the aromatic ring (data not shown). Tyrosine was modeled into this cavity, with the two possible catalytic mechanisms considered (Figures 3A and 3B). Using the conditions of the E_1cb mechanism, the amino group of tyrosine was positioned 1.5 Å from the methylidene carbon of MIO to simulate covalent attachment. From this position, the tyrosine hydroxyl was 2.4 Å from the His at position 89 (Figure 3A). Tyrosine was then modeled using the Friedel-Crafts-like mechanism with the same conditions described above, except the ortho carbon of the aromatic ring was

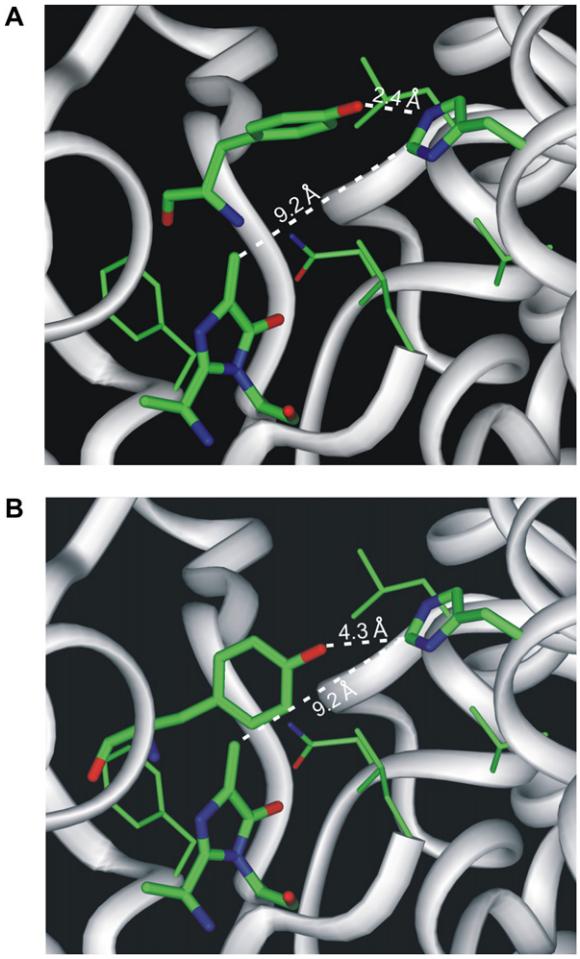

Figure 3. Investigation of the Active Site of the TAL Model

(A) The TAL model was created as described in Experimental Procedures. Using the model, an active site His residue (His89) was found which was not conserved across all PALs and TALs. The distance from the methylidene carbon of the MIO cofactor to His89 was found to be 9.2 Å. Also shown are conserved active site residues (Phe353 and Asn206) and residues forming the hydrophobic substrate binding pocket (Leu156, Leu202, and Val205). The substrate tyrosine was modeled into the active site using the proposed E_1cb mechanism. Under these conditions, the tyrosine hydroxyl group is approximately 2.4 Å away from His89.

(B) Tyrosine was modeled into the TAL active site using conditions described for the Friedel-Crafts-like acylation mechanism. Under these conditions, the substrate appears to occupy a similar relative position within the active site cavity but is moved away from His89, thereby increasing the distance from His89 to the tyrosine hydroxyl group to 4.3 Å.

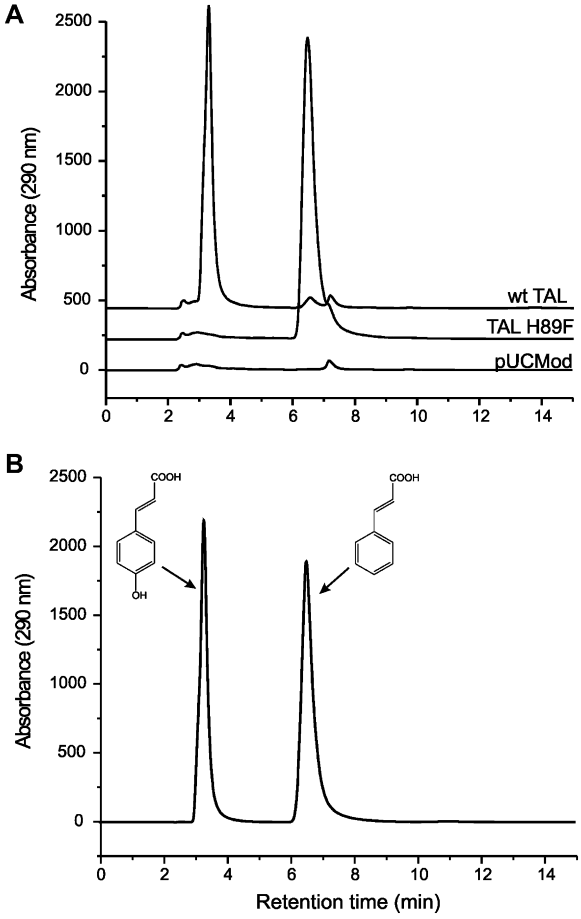

Figure 4. In Vivo Analysis of TAL and TAL H89F

(A) Cultures of *E. coli* expressing TAL, TAL H89F mutant, or pUCMod plasmid were extracted and run on HPLC, and chromatograms of each culture were overlaid for comparison. The bottom trace shows extraction of medium from *E. coli* pUCMod. The top trace shows extraction of medium from *E. coli* expressing TAL, showing a large peak at 3.2 min which corresponds to 4-coumaric acid. Also present is a small peak at 6.6 min corresponding to cinnamic acid. The middle trace shows extraction from TAL H89F culture, with the disappearance of the 4-coumaric acid peak and an increase in the peak corresponding to cinnamic acid.

(B) Standards of 4-coumaric acid (3.2 min) and cinnamic acid (6.6 min) were run under identical HPLC conditions for comparison of retention times.

modeled within 1.5 Å of MIO (Figure 3B). Using this mechanism, it appeared that the tyrosine hydroxyl was moved away from His89, now 4.3 Å away from that residue. This position would therefore appear to be poised for interaction with the substrate, and possibly have a role in substrate discrimination.

Kinetic Characterization of TAL H89F Mutant

To investigate the effect of substitution of His89 with Phe on substrate selection, a TAL H89F mutant was constructed and expressed in *E. coli*. Separate cultures of *E. coli* expressing wild-type TAL and *E. coli* with an empty plasmid were cultivated at the same time for comparison. The growth media of all cultures were extracted and analyzed by high-performance liquid chromatography (HPLC) for the presence of either 4-coumaric acid (TAL activity) or cinnamic acid (PAL activity). Shown in Figure 4A, no 4-coumaric or cinnamic acid was produced by *E. coli* containing an empty plasmid. In the culture extracts of *E. coli* expressing wild-type TAL, 4-coumaric acid was detected, along with trace amounts of cinnamic acid. *E. coli* culture expressing the H89F mutant showed complete loss of the peak corresponding to 4-coumaric acid (Figure 4A) and the appearance of a new major peak, which matched the retention time (Figure 4B) and UV/vis spectrum of cinnamic acid, indicating that the H89F mutant had completely lost TAL activity and acquired instead PAL activity in vivo.

Table 2. Steady-State Kinetic Parameters of TAL Mutants with Tyrosine and Phenylalanine as Substrates

TAL	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
Tyrosine			
WT	31.4 ± 8.8	3.4 ± 0.15	1.1×10^5
H89F	1541 ± 845.1	$0.0087 \pm 9.2 \times 10^{-4}$	5.6
H89Q	1488 ± 143.4	2.34 ± 0.05	1572
H89N	1399 ± 271.7	0.03 ± 0.01	21.4
H89A	ND	ND	ND
H89W	ND	ND	ND
Phenylalanine			
WT	5204 ± 1441	2.1 ± 0.45	403
H89F	269.1 ± 42.0	24.0 ± 0.089	8.9×10^4
H89Q	789.2 ± 87.8	26.2 ± 2.0	3.3×10^4
H89N	3456 ± 323	1.03 ± 0.07	298
H89A	4099 ± 652	0.18 ± 0.02	43.9
H89W	3777 ± 143	0.19 ± 0.01	50.3

Data represent the mean of six independent experiments ± standard deviation. ND, no activity detected with excess substrate and greater than 200 μg enzyme.

To confirm the apparent switch of substrate selectivity induced by the H89F substitution in vitro, the mutant enzyme was purified and its kinetic properties investigated. Activity measurements confirmed the pH optimum near 9 measured for TAL H89F was comparable to that found for the wild-type TAL (Figure 2C). TAL H89F was next assayed with tyrosine and found to have a 50-fold increase in K_m over wild-type TAL together with a nearly 400-fold drop in k_{cat} (Table 1). When TAL H89F was assayed with phenylalanine, a sharp increase in activity compared to wild-type was observed. The K_m for phenylalanine decreased 19-fold, while the k_{cat} for this substrate increased 10-fold. These corresponding k_{cat}/K_m values represent almost a reversal of those found for wild-type TAL with these same two substrates (Table 1).

Characterization of Additional TAL Mutants at Position 89

Additional mutants were constructed to investigate the influence of different amino acid residues at position 89 on TAL/PAL activity. As shown in Table 2, when His89 was changed, tyrosine binding was severely altered, as seen in the elevated K_m values for H89F, H89Q, and H89N. There is also an effect on turnover, as k_{cat} values for H89F and H89N show large decreases compared to wild-type. The H89Q mutant does not show this same sharp drop, suggesting that the size of the residue is important in enzyme turnover. Indeed, when His was replaced with Ala, no TAL activity was detected. Not surprisingly, the H89W mutation shows no TAL activity, most likely due to loss of the Tyr-OH/His interaction and the increased steric bulk associated with the Trp residue. For the substrate phenylalanine, PAL activity also appears to be dependent on the size of the amino acid (Table 2). Comparison of the K_m values for phenylalanine shows an increasing trend as the residues at position 89 decrease in size. However, the H89W mutant does not show increased PAL activity as would be expected, suggesting that perhaps the extra size associated with Trp excludes phenylalanine from proper interaction within the active site.

Replacement of Phe144 with His in *A. thaliana* PAL1

To determine whether the equivalent amino acid position was involved in substrate selection in PAL, and constituted a general PAL/TAL substrate selectivity switch, a PAL with low TAL activity was chosen for investigation. The selected enzyme, PAL1 from *Arabidopsis* (GenBank accession number: AAP59438), has recently been shown to have a very low k_{cat}/K_m (75 M^{-1} s^{-1}) for tyrosine together with a high k_{cat}/K_m (2.55×10^4 M^{-1} s^{-1}) for phenylalanine [20]. A mutant PAL was created in which Phe144 was replaced by His, and both wild-type PAL and PAL F144H were His tagged and purified from *E. coli* for in vitro characterization of their kinetic properties. Purified wild-type PAL and PAL F144H were then assayed as before with both phenylalanine and tyrosine (Table 1) (see Figure S1 in the Supplemental Data available with this article online). The wild-type enzyme had high catalytic efficiency with phenylalanine, together with a very low efficiency with tyrosine. The PAL F144H mutant, however, displayed a marked reduction (30-fold) in affinity for the substrate phenylalanine, and a 2-fold reduction in the turnover number for the reaction. This resulted in a k_{cat}/K_m value 80-fold lower than wild-type PAL for this same substrate. As anticipated, this drop in PAL activity was coupled with an increase in TAL activity. We observed a 2-fold drop in K_m and a 9-fold increase in k_{cat}, which corresponds to an 18-fold increase in the catalytic efficiency for tyrosine with the PAL F144H mutant. While PAL F144H kinetically remains a PAL, the increase in k_{cat}/K_m for tyrosine suggests that this residue has a similar function in PAL as it does in TAL, which is control of substrate selection.

Substrate Selectivity Switch as Predictive Tool for TAL or PAL Identification

Currently it is not possible to accurately predict whether an enzyme will have TAL, PAL or HAL activity based on sequence information alone. Inspection of the amino acid sequence of the described substrate selectivity switch region, however, may allow prediction of the likely substrate of this family of enzymes (Table S1). The known enzymes with HAL activity have Ser and His in the substrate selectivity region, while TALs and PALs have His or Phe at the switch position, followed by Leu (Figure 1B). One exception to this is the PAL from *R. toruloides*, which has a Gln at the position of Leu, which is preceded by a His at the switch position. This sequence conservation would therefore define a motif, which can be used to predict enzyme function based only upon sequence identity. In the case of HAL, the key motif would be Ser-His, and in the case of TAL and PAL activity, the motif would be His/Phe-Leu. The TAL/PAL activity can then be further refined by predicting that those enzymes that specifically show the His-Leu motif would be highly active with tyrosine.

To explore the possibility that the substrate selectivity switch could indeed be used as a predictive tool in identification of additional TALs or PALs, we performed a BLAST search with TAL from *Rba. sphaeroides* as the query sequence against microbial genome sequences in the NCBI database. Identified in this search was a protein from *Nostoc punctiforme* PCC 73102 annotated as a HAL (GenBank accession number: ZP_00105927). This enzyme showed 32% amino acid

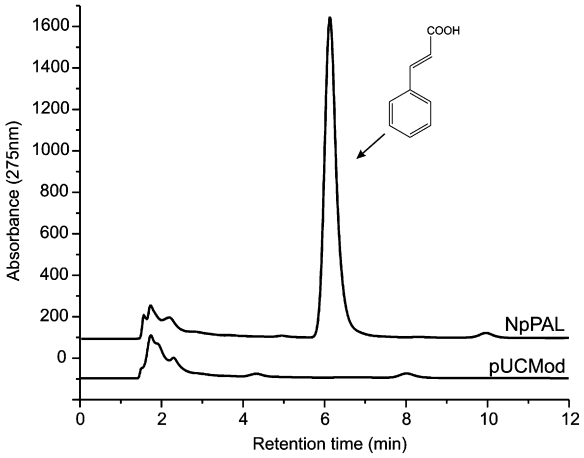

Figure 5. In Vivo Characterization of *N. punctiforme* PCC 73102 PAL

NpPAL was identified and cloned from *Nostoc* genomic DNA into pUCMod. *E. coli* expressing an empty pUCMod plasmid or NpPAL were grown overnight in modified M9 medium, extracted, and run on HPLC. HPLC chromatograms for each extraction are overlaid for comparison. The bottom trace shows extraction of medium from *E. coli* pUCMod culture. The top trace shows extraction of *E. coli* pUC-NpPAL culture and the appearance of a peak corresponding to cinnamic acid (6.6 min). No peaks corresponding to 4-coumaric acid or urocanic acid were detected by HPLC, indicating that NpPAL is a phenylalanine-specific PAL.

identity to TAL, a level commonly seen when comparing HALs and PALs from prokaryotes and eukaryotes alike [6, 26, 29] (Figure 1B). Comparison of its sequence with known aromatic amino acid lyases (Figure 1B) identified a Phe residue at the substrate selectivity switch position (Phe107), suggesting that this enzyme may actually be a PAL and not a HAL. To determine the activity of this enzyme, the predicted PAL (NpPAL) was cloned from *Nostoc* genomic DNA and expressed in *E. coli*. Culture media of *E. coli* pUC-NpPAL and control cultures of *E. coli* with empty pUCMod were extracted and run on HPLC to determine the in vivo activity of this enzyme. As seen in Figure 5, the *E. coli* culture expressing NpPAL had a new peak detected by HPLC not present in the culture of *E. coli* with empty pUCMod. This new peak had the identical retention time and UV/vis spectrum as a standard of cinnamic acid, indicating that NpPAL had PAL activity. No peak corresponding to 4-coumaric acid or urocanic acid was detectable, confirming the prediction that NpPAL is indeed a PAL and not a HAL or TAL.

Discussion

In this report, we describe the identification of a substrate selectivity switch in TAL, which, when mutated, converts the enzyme from primarily TAL activity to primarily PAL activity. The switch is a single amino acid (His89) identified in TAL through sequence alignment and modeling of the TAL structure. Rather than a simple expansion of the substrate range, this change converts the enzyme from a tyrosine-specific ammonia-lyase to a phenylalanine-specific ammonia-lyase. To the best of our knowledge, this result is a very rare example of a single amino acid dictating the substrate preference of an enzyme [30]. The structure of TAL has recently been solved (J.P. Noel, Howard Hughes Medical Institute and The Salk Institute for Biological Studies, La Jolla, CA, personal communication) and His89 was found to be in close proximity to 4-coumaric acid in the active site. Based on this finding, Noel and colleagues reasoned that this residue would be important for substrate selection, which nicely confirms our predictions made on sequence alignments and structural modeling of TAL.

The mechanism of how this amino acid position affects substrate selectivity is now apparent, considering the TAL model and kinetic data together. Across HALs, PALs, and TALs, there is strong conservation of active site residues determining the geometry of the active site (Figures 1B and 3). Among these are the invariant hydrophobic residues forming the binding pocket for the aromatic moiety of the substrate amino acid (TAL numbering: Leu86, Leu90, Leu156, Leu199, and Val202). Included within this group is the amino acid at position 89, identified as the substrate selection switch in this study. The presence of a His within this pocket creates a hydrogen bonding partner for the tyrosine hydroxyl group, and also interacts with the deamination product, 4-coumaric acid. In contrast, the presence of Phe would create an almost entirely hydrophobic pocket, and would strongly disfavor any possible interaction with a tyrosine hydroxyl group. Looking at catalytic efficiency (Table 1), it is obvious that having a His within the hydrophobic binding pocket does not necessarily prevent TAL or PAL activity. However, those enzymes having Phe within the hydrophobic pocket do appear to have a strong preference toward PAL activity. Looking at the ratios of catalytic efficiencies with the two substrates also illustrates this same effect. The two TALs have a 150- to 300-fold higher catalytic efficiency with tyrosine, whereas the PALs have an over 10,000-fold higher catalytic efficiency with phenylalanine. Even TAL H89F behaves just like a PAL, as its ratio is over 10,000 in favor of phenylalanine, suggesting that the presence of the active site Phe is strongly disfavoring an interaction with the substrate tyrosine, and further illustrating the importance of the His residue in TAL.

In addition to the importance of the active site His in hydrogen bonding, there appears to be an additional catalytic effect due to the size of the residue. It is clear from the kinetics, and TAL model structure, that the His interacts with the tyrosine hydroxyl. However, substitution of a similarly sized, hydrophilic residue (H89Q) results in an enzyme with reduced ability to bind tyrosine, but still able to catalyze deamination with near wild-type efficiency (Table 2). Substitution with a smaller hydrophilic residue (H89N) resulted in a similar K_m, but lowered the k_{cat} even further. This phenomenon is also evident when looking at PAL activity, which drops from maximum activity in the H89F mutant with each successively smaller amino acid substitution (Table 2). Exchange with a larger hydrophobic residue (H89W) results in an enzyme with very little activity, further demonstrating the importance of not only the side chain polarity/hydrophobicity but also the size of the amino acid at the switch position. Whether this feature is unique to the highly specific TALs or is generally true of all aromatic amino acid ammonia-lyases is currently under investigation.

Finally, using the switch rationale, we were able to predict the activity of an ammonia-lyase of unknown function. A PAL was cloned from *N. punctiforme* PCC 73102, which had previously been annotated as a HAL. This example demonstrates the importance of the amino acid residue at the switch location as a predictive tool. The discovery of this substrate selectivity switch should also increase the potential to develop ammonia-lyases as efficient biocatalysts [31, 32]. By tailoring the selectivity switch region, it may be possible to engineer these enzymes to deaminate a wider array of α-amino acids, or use other natural amino acids that have not previously been utilized. In doing so, it may lead to the alteration of products of the ammonia-lyase-containing pathways, such as changing the spectral properties of PYP or modulating the structure of various antibiotic compounds.

Significance

TAL is a recently described member of the aromatic amino acid ammonia-lyase family, producing the compound 4-coumaric acid from L-tyrosine, which is used in the biosynthesis of saccharomicin antibiotics as well as the chromophore of PYP. By comparison of the TAL amino acid sequence and model structure to closely related PALs and HALs, we identified a single His residue that influenced substrate specificity of this enzyme. When mutated, this position acted like a switch, effectively converting a highly selective TAL to a highly selective PAL, which is an unusual example of a complete alteration of substrate specificity by a single point mutation. The identified amino acid position was also found to be involved in substrate selection in PAL, further demonstrating the importance of this location in substrate selection across members of this enzyme family. Using the identity of the amino acid at the switch position as a guide, we were able to predict the activity of an unknown enzyme from *Nostoc punctiforme* PCC 73102 as a PAL and experimentally verify this prediction in vivo. This work has identified a critical amino acid in substrate selection in aromatic ammonia-lyases, which may enable the engineering of altered substrate specificities through simple point mutation, resulting in novel antibiotic structures or enhanced biocatalysts. Furthermore, using the identified sequence motif, it may now be possible to accurately predict the function of uncharacterized ammonia-lyases scattered in microbial genome databases, greatly expanding the number of known microbial aromatic ammonia-lyases.

Experimental Procedures

Strains and Culture Conditions

All cloning and DNA manipulation were carried out in *E. coli* JM109 by following standard techniques [33]. Cells were grown at 30°C with 250 rpm shaking in Luria Bertani (LB) medium [33] prior to plasmid harvest. Following sequencing, plasmids were transformed into *E. coli* strain BW27784 [34], provided by the *E. coli* Genetic Stock Center (New Haven, CT), and grown at room temperature (approximately 26°C) with 170 rpm shaking in Terrific Broth (TB) medium [33] for protein expression. For in vivo activity determination, *E. coli* BW27784 with the appropriate plasmid was grown in M9 medium modified by addition of 0.5% (v/v) glycerol and 1.25 g/L yeast extract. All media contained 100 μg/ml ampicillin or carbenicillin.

Expression Cloning

TAL was cloned previously from *Rba. sphaeroides* 158 genomic DNA [7]. An N-terminal 6× His tag was added to pUC-TAL by long-range polymerase chain reaction (PCR). Following PCR, template DNA was digested with DpnI (Promega, Madison, WI), and the 4000 bp fragment was gel purified, treated with T4 polynucleotide kinase (New England Biolabs, Ipswich, MA), and ligated with T4 DNA ligase (New England Biolabs) to make pUC-TAL_NHis. For construction of TAL mutants, site-directed mutagenesis was performed following the QuickChange protocol (Stratagene, La Jolla, CA) on the template pUC-TAL_NHis. *A. thaliana* PAL1 was cloned from an *A. thaliana* cDNA library purchased from the American Type Culture Collection (Manassas, VA; ATCC number 77500). The PCR product was restriction digested and ligated into the constitutive expression vector pUCMod [35] to make pUC-PAL. The PAL F144H mutation was constructed by QuickChange. Both pUC-PAL and pUC-PAL_F144H were N-terminal 6× His tagged using the same method as described above. NpPAL was identified by using the BLAST homology searching tool [36] with TAL from *Rba. sphaeroides* as the query sequence. PCR was done using genomic DNA from *Nostoc sp.* ATCC 29133 (ATCC number 29133D) and run under standard conditions. PCR products were restriction digested and ligated into pUCMod to create pUC-NpPAL.

Sequence Alignment and TAL Model Construction

Sequences were obtained from GenBank and aligned using the Multalin multiple sequence alignment tool [37]. The *Rba. sphaeroides* TAL model was created by submission of the TAL sequence to SwissModel [38, 39] using the alignment interface and using the A chain of the deposited *R. toruloides* PAL structure (PDB code: 1T6P) as the template. The MIO cofactor was built into the TAL model using the Schrödinger software suite (Schrödinger, New York, NY). The system was energy minimized for 100 steps using the steepest descent minimization algorithm, followed by 100 steps of conjugate gradient minimization. The final structure was energy minimized using optimized potentials for liquid simulations (OPLS) and surface generalized Born model. Substrate modeling and distance determinations were done using Insight II (Accelrys, San Diego, CA).

Tyrosine and Phenylalanine Ammonia-Lyase Activity In Vivo

Plasmids were transformed into *E. coli* BW27784 for determination of in vivo enzyme activity. Recombinant *E. coli* was grown overnight in 5 ml modified M9 medium, inoculated 1:100 into 50 ml modified M9 medium, and grown for an additional 24 hr. Expression levels of recombinant enzymes were checked by 12% SDS-PAGE using standard techniques [33]. For extraction of reaction products, 1 ml of culture was centrifuged at maximum speed (>13,000 rpm) to pellet cells. The supernatant was decanted to a fresh microfuge tube and the pH was adjusted to approximately 3.0 by addition of 50 μl 1 N hydrochloric acid, followed by overnight freezing at −20°C. Tubes were thawed and extracted twice with an equal volume (1 ml) of ethyl acetate. The pooled organic phase was dried under nitrogen gas and the residue was resuspended in 100 μl acetonitrile. All samples were stored at −20°C prior to HPLC analysis.

HPLC Analysis

Ten microliters of extract was applied to a Zorbax RX-C18 column (4.6 × 150 mm, 5 μm; Agilent Technologies, Palo Alto, CA) with an additional SB-C18 analytical guard column (4.6 × 12.5 mm, 5 μm; Agilent) using an Agilent 1100 HPLC system equipped with a photodiode array detector. Elution was performed with an isocratic mobile phase of acetic acid:water:acetonitrile (0.7:69.3:30) and a constant flow rate of 1.0 ml/min. Compound peaks were identified by comparison to retention times and UV/vis spectra of standard compounds.

Protein Purification

E. coli expressing 6× His-tagged protein was grown overnight in 5 ml TB medium at 30°C. Following growth overnight, 1 liter TB medium was inoculated 1:200 with preculture and grown at room temperature with 170 rpm shaking. After 24 hr growth, cells were pelleted by centrifugation at 4000 × g and washed with 50 ml wash buffer (50 mM sodium phosphate, 300 mM NaCl [pH 8.0]). Cells were resuspended in wash buffer on ice and lysed by sonication,

followed by centrifugation at 10,000 × g for 30 min to pellet cellular debris. Cleared supernatant was loaded onto a packed BD TALON (Clontech, Mountain View, CA) gravity column, and eluted with wash buffer containing 150 mM imidazole. Following elution, protein was precipitated overnight on ice in 70%–80% ammonium sulfate and collected by centrifugation at 4000 × g for 30 min at 4°C. The protein pellet was resuspended, and soluble protein was loaded onto a PD-10 desalting column (Amersham, Piscataway, NJ) equilibrated with 25 ml 50 mM sodium phosphate buffer (pH 8.0). Following desalting, protein was quantified using a BCA protein assay kit (Pierce, Rockford, IL).

Enzyme Assays

PAL activity was determined by monitoring cinnamic acid formation at 275 nm ($\varepsilon = 15{,}560\ M^{-1}\ cm^{-1}$) in 1 cm quartz cuvettes. TAL activity was determined by detecting 4-coumaric acid at 310 nm ($\varepsilon = 9{,}554\ M^{-1}\ cm^{-1}$). Substrate concentration for PAL assays was varied between 20 mM and 5 μM, and for TAL assays between 5 mM and 5 μM. Enzyme concentration used in assays varied between 10 and 25 μg for wild-type enzyme and 100 and 200 μg for less active mutants. pH optima were determined between the values of 6.5 and 11.0 using 100 mM bis-Tris propane (pH 6.5–9.0) and 100 mM N-cyclohexyl-3-aminopropanesulfonic acid (CAPS) (pH 9.0–11.0). All subsequent assays were carried out in a final volume of 1 ml with 100 mM bis-Tris propane (pH 9.0) or 100 mM 2-amino-2-methyl-1,3-propanediol (pH 9.0) at 37°C. To determine the influence of metals on enzyme activity, assays were performed in 100 mM Tris-HCl (pH 8.0) with saturating substrate concentration and 0.01–1 mM $ZnSO_4$, $MgSO_4$, $MnCl_2$, or $CaCl_2$, or 50–500 mM EDTA (pH 8.0).

Supplemental Data

Supplemental Data include one figure and one table and can be found with this article online at http://www.chembiol.com/cgi/content/full/13/12/1317/DC1/.

Acknowledgments

The authors gratefully acknowledge Dr. Gordon Louie and Dr. Joseph Noel (Salk Institute) for the communication of their TAL structural data immediately prior to publication. The authors also acknowledge Dr. Yuk Yin Sham for assistance with TAL modeling. This investigation has been supported by a grant from the David and Lucile Packard Foundation (grant 2001-18996). K.T.W. was supported by an NIGMS/NIH Biotechnology Training Grant (T32 GM08347) and the University of Minnesota Presidential Initiative on Biocatalysis.

Received: August 10, 2006
Revised: October 5, 2006
Accepted: October 17, 2006
Published: December 22, 2006

References

1. Poppe, L., and Retey, J. (2005). Friedel-Crafts-type mechanism for the enzymatic elimination of ammonia from histidine and phenylalanine. Angew. Chem. Int. Ed. Engl. *44*, 3668–3688.
2. Winkel-Shirley, B. (2001). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol. *126*, 485–493.
3. Williams, J.S., Thomas, M., and Clarke, D.J. (2005). The gene stlA encodes a phenylalanine ammonia-lyase that is involved in the production of a stilbene antibiotic in *Photorhabdus luminescens* TT01. Microbiology *151*, 2543–2550.
4. Xiang, L., and Moore, B.S. (2002). Inactivation, complementation, and heterologous expression of encP, a novel bacterial phenylalanine ammonia-lyase gene. J. Biol. Chem. *277*, 32505–32509.
5. Sarkissian, C.N., and Gamez, A. (2005). Phenylalanine ammonia lyase, enzyme substitution therapy for phenylketonuria, where are we now? Mol. Genet. Metab. *86* (*Suppl. 1*), S22–S26.
6. Kyndt, J.A., Meyer, T.E., Cusanovich, M.A., and Van Beeumen, J.J. (2002). Characterization of a bacterial tyrosine ammonia lyase, a biosynthetic enzyme for the photoactive yellow protein. FEBS Lett. *512*, 240–244.
7. Watts, K.T., Lee, P.C., and Schmidt-Dannert, C. (2004). Exploring recombinant flavonoid biosynthesis in metabolically engineered *Escherichia coli*. ChemBioChem *5*, 500–507.
8. Berner, M., Krug, D., Bihlmaier, C., Vente, A., Muller, R., and Bechthold, A. (2006). Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete *Saccharothrix espanaensis*. J. Bacteriol. *188*, 2666–2673.
9. Cusanovich, M.A., and Meyer, T.E. (2003). Photoactive yellow protein: a prototypic PAS domain sensory protein and development of a common signaling mechanism. Biochemistry *42*, 4759–4770.
10. Baedeker, M., and Schulz, G.E. (2002). Autocatalytic peptide cyclization during chain folding of histidine ammonia-lyase. Structure *10*, 61–67.
11. Schulz, W., Eiben, H.G., and Hahlbrock, K. (1989). Expression in *Escherichia coli* of catalytically active phenylalanine ammonia-lyase from parsley. FEBS Lett. *258*, 335–338.
12. Schwede, T.F., Retey, J., and Schulz, G.E. (1999). Crystal structure of histidine ammonia-lyase revealing a novel polypeptide modification as the catalytic electrophile. Biochemistry *38*, 5355–5361.
13. Shi, W., Dunbar, J., Jayasekera, M.M., Viola, R.E., and Farber, G.K. (1997). The structure of L-aspartate ammonia-lyase from *Escherichia coli*. Biochemistry *36*, 9136–9144.
14. Calabrese, J.C., Jordan, D.B., Boodhoo, A., Sariaslani, S., and Vannelli, T. (2004). Crystal structure of phenylalanine ammonia lyase: multiple helix dipoles implicated in catalysis. Biochemistry *43*, 11403–11416.
15. Ritter, H., and Schulz, G.E. (2004). Structural basis for the entrance into the phenylpropanoid metabolism catalyzed by phenylalanine ammonia-lyase. Plant Cell *16*, 3426–3436.
16. Schuster, B., and Retey, J. (1995). The mechanism of action of phenylalanine ammonia-lyase: the role of prosthetic dehydroalanine. Proc. Natl. Acad. Sci. USA *92*, 8433–8437.
17. Hanson, K.R., and Havir, E.A. (1970). L-phenylalanine ammonia-lyase. IV. Evidence that the prosthetic group contains a dehydroalanyl residue and mechanism of action. Arch. Biochem. Biophys. *141*, 1–17.
18. Hanson, K.R., and Havir, E.A. (1981). In The Biochemistry of Plants, *Volume 7*, P.K. Stumpf and E.E. Conn, eds. (New York: Academic Press).
19. Appert, C., Logemann, E., Hahlbrock, K., Schmid, J., and Amrhein, N. (1994). Structural and catalytic properties of the four phenylalanine ammonia-lyase isoenzymes from parsley (*Petroselinum crispum* Nym.). Eur. J. Biochem. *225*, 491–499.
20. Cochrane, F.C., Davin, L.B., and Lewis, N.G. (2004). The *Arabidopsis* phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. Phytochemistry *65*, 1557–1564.
21. Rosler, J., Krekel, F., Amrhein, N., and Schmid, J. (1997). Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. Plant Physiol. *113*, 175–179.
22. Neish, A.C. (1961). Formation of m- and p-coumaric acids by enzymatic deamination of the corresponding isomers of tyrosine. Phytochemistry *1*, 1–24.
23. Breinig, S., Qi, W.W., Sariaslani, F.S., Vannelli, T.M., and Xue, Z. (October 2005). DNA and amino acid sequences of a tyrosine-inducible tyrosine ammonia lyase enzyme from the yeast *Trichosporon cutaneum*. U.S. patent 6,951,751.
24. Sawada, S., Kumagai, H., Yamada, H., Hill, R.K., Mugibayashi, Y., and Ogata, K. (1973). Stereochemistry of ammonia elimination from L-tyrosine with L-phenylalanine ammonia-lyase. Biochim. Biophys. Acta *315*, 204–207.
25. Hodgins, D.S. (1971). Yeast phenylalanine ammonia-lyase. Purification, properties, and the identification of catalytically essential dehydroalanine. J. Biol. Chem. *246*, 2977–2985.
26. Xiang, L., and Moore, B.S. (2005). Biochemical characterization of a prokaryotic phenylalanine ammonia lyase. J. Bacteriol. *187*, 4286–4289.
27. Klee, C.B. (1972). Metal activation of histidine ammonia-lyase. Metal ion-sulfhydryl group relationship. J. Biol. Chem. *247*, 1398–1406.
28. Rother, D., Poppe, L., Viergutz, S., Langer, B., and Retey, J. (2001). Characterization of the active site of histidine

ammonia-lyase from *Pseudomonas putida*. Eur. J. Biochem. *268*, 6011–6019.
29. Rother, D., Poppe, L., Morlock, G., Viergutz, S., and Retey, J. (2002). An active site homology model of phenylalanine ammonia-lyase from *Petroselinum crispum*. Eur. J. Biochem. *269*, 3065–3075.
30. Johnson, E.T., Ryu, S., Yi, H., Shin, B., Cheong, H., and Choi, G. (2001). Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. Plant J. *25*, 325–333.
31. Paizs, C., Katona, A., and Rétey, J. (2006). Chemoenzymatic one-pot synthesis of enantiopure L-arylalanines from arylaldehydes. Eur. J. Org. Chem. *2006*, 1113–1116.
32. Paizs, C., Katona, A., and Rétey, J. (2006). The interaction of heteroaryl-acrylates and alanines with phenylalanine ammonia-lyase from parsley. Chemistry *12*, 2739–2744.
33. Sambrook, J. (2001). Molecular Cloning: A Laboratory Manual, *Volume 3*, Third Edition (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
34. Khlebnikov, A., Datsenko, K.A., Skaug, T., Wanner, B.L., and Keasling, J.D. (2001). Homogeneous expression of the P(BAD) promoter in *Escherichia coli* by constitutive expression of the low-affinity high-capacity AraE transporter. Microbiology *147*, 3241–3247.
35. Schmidt-Dannert, C., Umeno, D., and Arnold, F.H. (2000). Molecular breeding of carotenoid biosynthetic pathways. Nat. Biotechnol. *18*, 750–753.
36. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.
37. Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. *16*, 10881–10890.
38. Guex, N., and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis *18*, 2714–2723.
39. Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2003). SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. *31*, 3381–3385.