Crystal Structure of Phenylalanine Ammonia Lyase: Multiple Helix Dipoles Implicated in Catalysis[†,‡]

Joseph C. Calabrese,* Douglas B. Jordan,*[,#] Amechand Boodhoo,[§] Sima Sariaslani, and Todd Vannelli[||]

DuPont Central Research and Development, Experimental Station, Wilmington, Delaware 19880-0228

Received May 10, 2004; Revised Manuscript Received July 7, 2004

ABSTRACT: The first three-dimensional structure of phenylalanine ammonia lyase (PAL) has been determined at 2.1 Å resolution for PAL from *Rhodosporidium toruloides*. The enzyme is structurally similar to the mechanistically related histidine ammonia lyase (HAL), with PAL having an additional ∼160 residues extending from the common fold. We propose that catalysis (including lowering the pK_a of nonacidic C3 of L-phenylalanine for an E1cb mechanism) is potentially governed by dipole moments of seven α helices associated with the PAL active site (six positive poles and one negative pole). Cofactor 3,5-dihydro-5-methylidene-4*H*-imidazol-4-one (MIO) resides atop the positive poles of three helices, for increasing its electrophilicity. The helix dipoles appear fully compatible with a model of phenylalanine docked in the active site of PAL having the first covalent bond formed between the amino group of substrate and the methylidene group of MIO: 12 highly conserved residues (near the N termini of helices for enhancing function) are poised to serve roles in substrate recognition, MIO activation, product separation, proton donation, or polarizing electrons from the phenyl ring of substrate for activation of C3; and a highly conserved His residue (near the C terminus of the one helix that directs its negative pole toward the active site to increase the residue's basicity) is positioned to act as a general base, abstracting the pro-*S* hydrogen from C3 of substrate. A similar mechanism is proposed for HAL, which has a similar disposition of seven α helices and similar active-site residues. The helix dipoles appear incompatible with a proposed mechanism that invokes a carbocation intermediate.

Phenylalanine ammonia lyase (PAL,[1] EC 4.3.1.5) catalyzes the nonoxidative deamination of L-phenylalanine to *trans*-cinnamic acid and ammonia, thus facilitating the commitment step in phenylpropanoid pathways that produce lignins, coumarins, and flavonoids in plants, fungi, and bacteria (*1*). Depending on the enzyme's biological source, it may act with greater or lesser efficiency on tyrosine, with the variation among relative substrate specificities for phenylalanine versus tyrosine documented to span a factor of 1 600 000 (*2*, *3*). At one extreme, PAL from dicotyledonous plants is highly specific for phenylalanine, and at the other extreme the enzyme has greater specificity for tyrosine in certain photosynthetic bacteria where it serves to supply *p*-hydroxycinnamic acid as the chromophore for the photoactive yellow protein; the latter enzyme has properly been termed a tyrosine ammonia lyase (*3*). In monocotyledon plants, PAL has similar substrate specificities for tyrosine and phenylalanine and the catalyst serves double duty. Having roles of mediating the first step of the biosynthesis of lignins for forming lignocelluloses, the second most abundant biopolymer following cellulose, and that of more rare natural products, regulation of PAL activity has been of some interest. Industrial applications include engineering of PAL to improve its activity toward tyrosine for producing *p*-hydroxycinnamic acid (*4*).

The family of L-amino acid ammonia lyases mediate the abstraction of a proton from the β carbon and the removal of the amino group from their substrate amino acids, and the individual enzymes employ different strategies in doing so (*5*, *6*): for example, the reactions catalyzed by aspartate ammonia lyase (AAL) and methylaspartate ammonia lyase (MAL) have been shown to eliminate ammonia from their substrates through carbanion intermediates without the use of an organic cofactor (*7*, *8*). In the E1cb-like mechanisms, it is thought that the γ carboxyl group of substrate activates the β carbon to lose a proton and form an aci-carboxylate intermediate, followed by the elimination of the amino group as ammonia. Structurally, AAL belongs to the fumarase superfamily and MAL to the enolase superfamily. Since the β methylene groups adjacent to the aromatic rings of

[†] This research was carried out in part at the National Synchrotron Light Source, Brookhaven National Laboratory, which is supported by the Division of Materials Sciences and Division of Chemical Sciences, U.S. Department of Energy, under Contract DE-AC02-98CH10886.

[‡] Atomic coordinates and structure factors for the two PAL structures have been deposited in the Protein Data Bank (http://www.rcsb.org) under Accession Numbers 1T6J and 1T6P.

* To whom correspondence should be addressed: e-mail joseph.c.calabrese@usa.dupont.com (J.C.C.) or jordand@ncaur.usda.gov (D.B.J.).

[#] Present address: National Center for Agricultural Utilization Research, 1815 N. University St., Peoria, IL 61604.

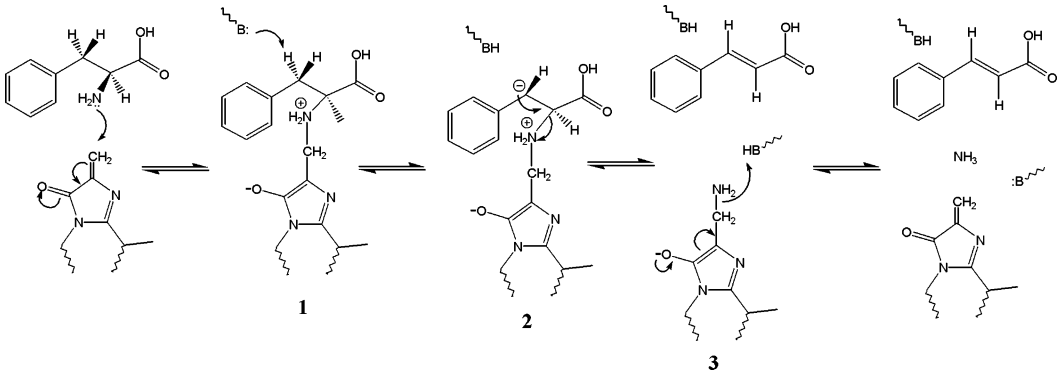
[§] Present address: Eli Lilly and Co, Lilly Corporate Center, Indianapolis, IN 46285.

[||] Present address: Ludwig Institute for Cancer Research, 307 Stocking Hall, Cornell University, Ithaca, NY 14853.

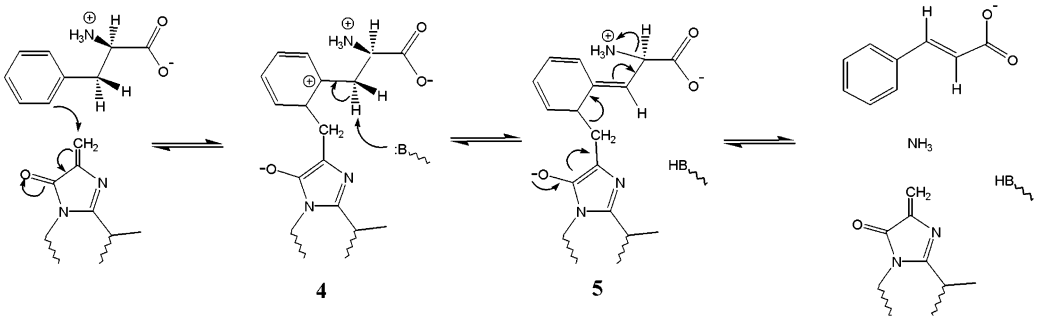
[1] Abbreviations: MIO, 3,5-dihydro-5-methylidene-4*H*-imidazol-4-one; E1cb, elimination first-order conjugate base; PAL, phenylalanine ammonia lyase; HAL, histidine ammonia lyase; AAL, aspartate ammonia lyase; MAL, methyl aspartate ammonia lyase; TAL, tyrosine ammonia lyase; TAM, tyrosine aminomutase; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid.

10.1021/bi049053+ CCC: $27.50
Published on Web 08/20/2004

Scheme 1



Scheme 2



L-phenylalanine and L-histidine are less acidic than that of L-aspartate, there has been considerable interest in and numerous proposals for the mechanisms of the amino acid lyases acting on these substrates (*9*−*13*). Recently, X-ray structures of histidine ammonia lyase (HAL) have shown that it contains the cofactor 3,5- dihydro-5-methylidene-4*H*-imidazol-4-one (MIO) (*14*−*16*), which is formed by cyclization and dehydration of residues within its conserved Ala-Ser-Gly sequence (*17*, *18*). MIO is a more electrophilic version of dehydroalanine that was previously thought to serve as cofactor in PAL and HAL. PAL enzymes have the same Ala-Ser-Gly signature as HAL and have been deduced to contain MIO by other methods (*17*, *18*). In PAL, kinetic isotope effect (KIE) studies have determined that the strength of MIO as an electrophile is insufficient for reacting with and removing the amino group of alternate substrate dihydrophenylalanine followed by loss of the C3 hydrogen as in an E1 mechanism; instead the KIE results point to the opposite end of the kinetic spectrum, an E1cb-like mechanism, where the C3 hydrogen is removed first, followed by the elimination of ammonia (*20*). Such a thorough kinetic study has not been reported for HAL.

Two possible mechanisms for PAL and HAL, relevant to current considerations, are shown for the PAL reaction in Schemes 1 and 2 (*1*, *9*−*20*). Both are E1cb-like. In the steps of Scheme 1, the methylidene group of MIO reacts with the NH_2 group of substrate forming **1**; the electron-withdrawing capability of the substrate's aromatic side chain and the positively charged nitrogen of **1** serve to activate C3 of substrate for deprotonation by an enzyme general base, generating carbanion **2**; the bond between C2 of substrate and the amino group is broken, forming **3** and product cinnamate; and the MIO−NH_2 complex **3** is protonated by a general acid to eliminate NH_3 and regenerate MIO. Reservations on this mechanism center on the ability to promote development and stabilization of carbanion intermediate **2**, and the mechanism of Scheme 2 was devised, in part, to provide a stronger means of supporting an E1cb mechanism. In the steps of Scheme 2, in a Friedel−Crafts-type reaction the aromatic ring of phenylalanine attacks the MIO electrophile to form carbocation **4**, which would stabilize intermediate **5** formed by removal of the substrate's C3 hydrogen. Collapse of the system to product occurs with the elimination of NH_3 and the release of cinnamate from the MIO cofactor. Reservations on this mechanism center on the potentially large energy barrier that must be surpassed in forming the carbocation intermediate. Biochemical support for the outlined mechanisms of Schemes 1 and 2 stems from alternate substrate studies for PAL (*21*), solvent isotope exchange studies in HAL (*22*), and KIE studies on the perdeuterated and ortho-tritiated phenyl ring of phenylalanine for PAL (*21*, *23*). Both mechanisms can be reconciled with the alternate substrate, the solvent isotope exchange, and the perdeuterated phenyl ring KIE studies. Even the inverse secondary KIE result (on *V*/*K*) with ortho-tritiated phenylalanine, which would seem to strongly support Scheme 2 over Scheme 1, can be reconciled with Scheme 1 if the observed KIE were of similar magnitude as an inverse isotope effect on binding, as has been found for hydrogen-bonding interactions (*24*) and hydrophobic interactions (*25*) in other systems; the tritium binding effect in PAL remains undetermined. Thus, the opposed mechanisms described by Schemes 1 and 2 seem to resist biochemical attempts to resolve the conundrum.

The unliganded structure of HAL has been used to model substrate histidine into the active site and react with MIO according to mechanism of Scheme 2 (*14*, *16*), and one of the reports briefly considered and dismissed the possibility

of the mechanism of Scheme 1 (*16*). In addition, a homology model of PAL, based on the HAL structure, has been used to dock the substrate phenylalanine into the active site according to the mechanism of Scheme 2 (*26*). As such, the mechanism of Scheme 2 might be considered "generally accepted" in the recent literature. In this work, we report two X-ray structures of PAL determined from crystals grown in the presence of its reaction products, *trans*-cinnamate and ammonia. This constitutes the first revelation of the three-dimensional structure of the enzyme. Indeed, there is an MIO cofactor residing in PAL, here observed containing a putative NH_2 adduct. The structural results are interpreted in terms of Schemes 1 and 2, respecting the conserved residues of PAL, the activities of site-directed mutations, and other determinations relating to the PAL and HAL mechanisms. The structures were determined for PAL from *Rhodosporidium toruloides* (also known as *Rhodotorula glutinis* and otherwise depending on the organism's stage in the life cycle), on which many of the relevant kinetic studies have been performed.

EXPERIMENTAL PROCEDURES

Expression and Purification. Escherichia coli cells, B834(DE3)pLysS, containing plasmid pEAL (PAL inserted into pET-24a) were grown overnight at 37 °C in 50 mL of minimal medium containing methionine, 50 μg/mL kanamycin, and 34 μg/mL chloramphenicol (*27*). Cells were harvested, suspended in 1 L of the minimal medium containing Se-methionine, and grown at 37 °C. When the cell density reached an OD_{600} of 0.5−1, the culture was brought to 0.1 mM isopropyl β-D-thiogalactoside. Cells were grown overnight at ∼23 °C, reaching an OD_{600} of 2.9, and harvested by centrifugation.

Protein isolation steps were performed at 4 °C. Cells were washed with 100 mM Tris-HCl, pH 8.5, and suspended to 1 g of cells/mL of buffer, containing 100 mM Tris-HCl, pH 8.5, 20 mM DTT, 1 mM EDTA, a Complete Mini tablet (Roche, no. 1836153), leupeptin (1 μg/mL), pepstatin A (1 μg/mL), bestatin (40 μg/mL), and DNase (1 μg/mL, Sigma D4527). Cells were passed twice through a French press at ∼18 000 psi, and cell debris was removed by centrifugation. The protein was purified by three chromatography steps: (1) anion-exchange chromatography (20 × 165 mm, 50 μm HQ column; Amersham Biosystems) with a gradient ranging from 5 mM Tris-HCl, pH 8.5/10 mM DTT/0.2 mM EDTA to 0.5 M NaCl/10 mM Tris-HCl, pH 8.5/5 mM DTT/0.2 mM EDTA; (2) hydrophobic interaction chromatography (20 mm × 167 mm, 50 μM PE column; Poros) with a gradient ranging from 1 M $(NH_4)_2SO_4$/10 mM Tris-HCl, pH 8.5/5 mM DTT/1 mM EDTA to 10 mM Tris-HCl, pH 8.5/5 mM DTT/1 mM EDTA; and (3) gel-filtration chromatography (16 × 60 mm Superdex 200 column; Amersham Biosystems), equilibrated in 50 mM Tris-HCl, 0.1 M NaCl, and 5 mM DTT, pH 7.5. The protein was concentrated to 8 mg/mL by use of a Centricon system (Amersham Biosystems). SDS−PAGE analysis of the preparation indicated that it was predominately PAL with some contaminating bands.

Crystallization. Crystallizations were performed by the hanging-drop vapor diffusion method, adding 2 μL of the protein solution (8 mg/mL supplemented with 1 mM *trans*-cinnamate) to 2 μL of the well solution. Crystals, used for analysis, were obtained by macro seeding of 1-day-old equilibrated drops, and they appeared fully grown following 1−2 days of incubation at room temperature. The reservoir solution that produced the monoclinic crystal contained 17−18% poly(ethylene glycol) (PEG) 3000 and 0.1 M citrate, pH 5.6; that for the trigonal crystal contained the above plus 0.5% octyl β-glucoside.

Data Collection, Structure Determination, and Refinement. Data for the trigonal crystal were collected at the National Synchrotron Light Source of Brookhaven National Laboratories at line X12C, and data for the monoclinic crystal were collected at the DND beam line at the Advanced Photon Source of Argonne National Laboratories. Data were processed with DENZO and Scalepack (*28*). Direct methods (*29*, *30*) applied to the anomalous difference data yielded 54 of 60 required Se positions in space group $P3_1$ for the trigonal form. Examination of this solution revealed that the Se distribution was consistent with either of two space groups of the higher Laue point group symmetry $P3_12$. Phases were refined with SHARP (*31*) in both subgroups, and the correct handedness was established by a comparison of the respective Fourier maps to be consistent with space group $P3_221$. The symmetry of this space group relates a single tetramer via its diagonal crystallographic 2-fold axis, requiring model building of two independent polymeric chains. The monoclinic structure was solved via molecular replacement by use of the program MOLREP (*32*) and coordinates for the tetramer generated from the trigonal structure. Consistent with the cell volume, the best solution was obtained for two tetramers in the asymmetric unit with a correlation coefficient of 0.41. The model required building two independent tetrameric units (eight polymeric chains containing 5728 residues). Electron density maps were displayed and models were manually constructed with the programs O (*33*) and XtalView/Xfit (*34*).

Refinement cycles were performed with the program CNX (Accelyris, Inc.), a commercial version of CNS (*35*), with coordinates for the MIO moiety initially taken directly from the HAL structure (*14*). The standard CNX parameter and topology files were modified to incorporate the MIO moiety as an integral part of the polypeptide backbone. Data (2.5%) were reserved for the calculation of R_{free} (*36*). Alternate cycles of simulated annealing, omit maps, and the addition of water molecules were continued until there was no significant improvement in R_{free} values. Careful inspection at latter stages of refinement revealed that all 10 independent active sites of the two data sets contained an extra peak near the methylidene group of MIO. This was further refined as the nitrogen of the reaction product, presuming $(NH_4)_2SO_4$ carryover from protein purification. Final refinements were continued with a MIO−NH_2 rigid group built via Sybyl 6.7 (Tripos, Inc.). Amide groups were reoriented to yield the most compatible hydrogen bonding by use of the program Reduce version 2.21 (*37*).

The program TOP was used for making a comparison of HAL and PAL tertiary structures (*38*). With the exception of the above-mentioned programs, incidental crystallographic calculations (Fourier maps, density modification, etc.) were carried out with the CCP4 program suite (*39*). Color figures were generated by use of Ribbons (*40*), and sequence alignments were performed through CLUSTAL W (*41*).

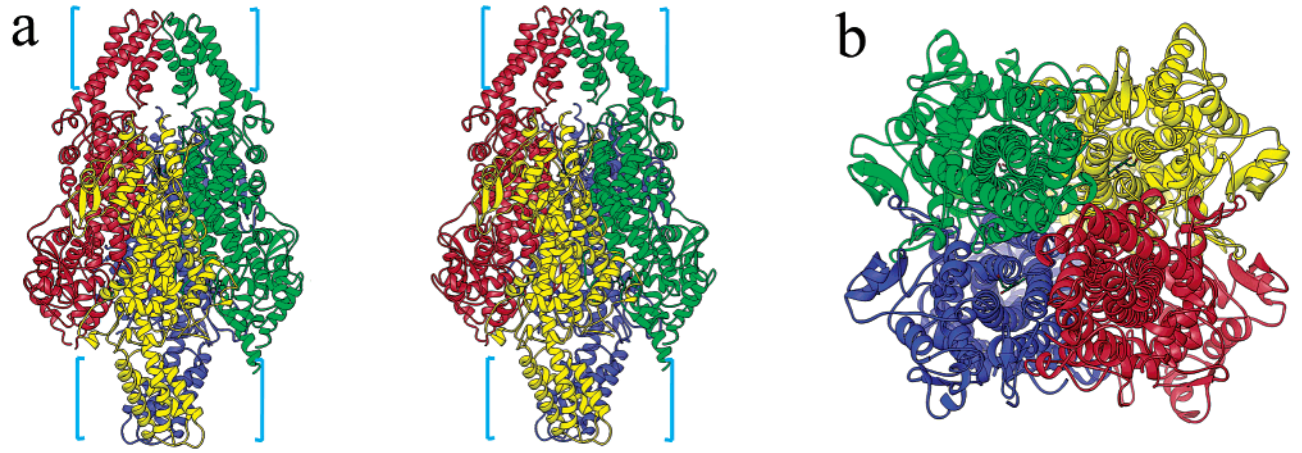


FIGURE 1: Quaternary structure of PAL. The four individual monomers are color-coded in red, green, blue, and yellow, displaying the approximate 222 symmetry of the PAL tetramer. (a) Stereoview of PAL from a side perspective. The bracketed areas represent the residues, arranged in a fan, that are present in PAL and absent in HAL. (b) Top perspective of the PAL tetramer (90° offset from that of panel a). This view looks down into the active sites of two of the subunits; the other two active sites would be seen from a bottom view.

Table 1: Data Collection and Refinement Statistics

crystal form	trigonal[a]	monoclinic[a]
PDB code	1T6J	1T6P
beam line	NSLS 12C	APS DND 5ID
wavelength (Å)	0.978 85,[b] 0.978 66,[b,c] 0.963 50[b]	0.980 65[c]
space group	$P3_221$	$P2_1$
cell dimensions *a*, *b*, *c* (Å)	107.9, 107.9, 204.4	105.2, 181.0, 149.2
unique β angle (deg)		96.95
resolution of the structure (Å)	25.0−2.10	25.1−2.70
unique reflections	75 419	146 876
completeness (%)	90.0[d]	95.3[d]
redundancy	6.33	3.25
$I/\sigma(I)$	20.5	18.7
R_{merge} (%)	8.9	7.1
no. of Se sites	30	120
monomers per asymmetric unit	2	8
solvent content (%)	45.6	46.8
amino acid residues	1294	5219
no. of non-hydrogen atoms	10 300	40 380
water molecules	411	651
R_{work}/R_{free}	0.251/0.299	0.194/0.253
rmsd of bond lengths (Å)/ bond angles (deg) from ideality	0.0055/1.26	0.0067/1.37
avg *B* of main-chain atoms (Å^2)	49.4	37.9
avg *B* of side-chain atoms (Å^2)	50.2	39.5

[a] Data were collected on a single crystal cooled to −180 °C. [b] Wavelength used for MAD phasing. [c] Wavelength used for refinement. [d] The highest-resolution shell of the trigonal data set (2.14−2.10 Å) is 49% complete; that of the monoclinic data set (2.75−2.70 Å) is 80.6% complete.

RESULTS AND DISCUSSION

PAL Structure. Two X-ray structures of PAL, crystallized in the presence of *trans*-cinnamate, were determined by using MAD phasing and Se-Met substituted protein. Diffraction data for the trigonal form, containing two subunits per unit cell, extended to 2.1 Å, and that for the monoclinic form, containing eight subunits per unit cell, extended to 2.7 Å. Statistics for data collection and refinement are summarized in Table 1. The trigonal data was affected by an imperfect superposition of two orientations of molecules in the lattice, resulting in a weak superlattice along the *c* axis and doubling of the data along this direction. Since the average intensity of superlattice reflections was less than 1% that of the sublattice data, the superlattice data were excluded from the structure solution and refinement. The superposition is evidenced primarily near the disordered N-terminal region, where the backbone trace diverges into two directions. The incomplete superposition likely accounts for the higher than expected R_{work} and R_{free} values for a resolution of this order. It was considered necessary to refine the lower-resolution, monoclinic data set because the weak superlattice of the trigonal data set could potentially increase the number of residues interpreted as disordered. However, comparison of the disordered regions between the data sets indicates that that they are similar: the monoclinic form lacked electron density for residues 1−38, 105−121, and 349−358 for each of the eight polypeptides in the asymmetric unit, while the trigonal form lacked electron density for residues 1−39, 105−126, and 349−358 for its two polypeptides. Common to all 10 subunits refined are residues 40−104, 127−344, and 361−716. Among the 10 crystallographically independent subunits, the root-mean-square deviation from fitting common Cα atoms is 0.18 Å.

PAL from *R. toruloides* is a homotetramer (716 residues per subunit; molecular mass 76 880 Da) arranged in 222 symmetry (Figures 1 and 2). Each subunit assumes a "seahorse" shape interlocking head-to-tail with two other subunits, maximizing adjacent subunit interactions and yielding a close-fitting tetramer. Formation of the tetramer buries a surface area of 66 386 Å^2 among the monomers or 58% of their combined surfaces. Of 66 interactions between adjacent subunits, 25 involve hydrogen bonds between Asp and Glu carboxylate oxygens and NH_2 and OH moieties, and there is a prominent band of Asp and Glu interactions with Arg side chains between adjacent subunits in the vicinity of the central bundle of helices. As well, the tetramer assembly leads to a cluster of four vicinal cysteines (residues 140, 455, 467, and 530) with the sulfur atoms of Cys467 and Cys530 separated by 3.62 Å. Presumably, the presence of DTT in buffers prevented the formation of disulfides. Regulatory mechanisms involving disulfides of PAL, if any, are unknown.

The structure of the main body of PAL has a similar fold as the HAL structure (*14*), with a root-mean-square deviation of 1.4 Å from fitting the 440 closest Cα pairs between HAL

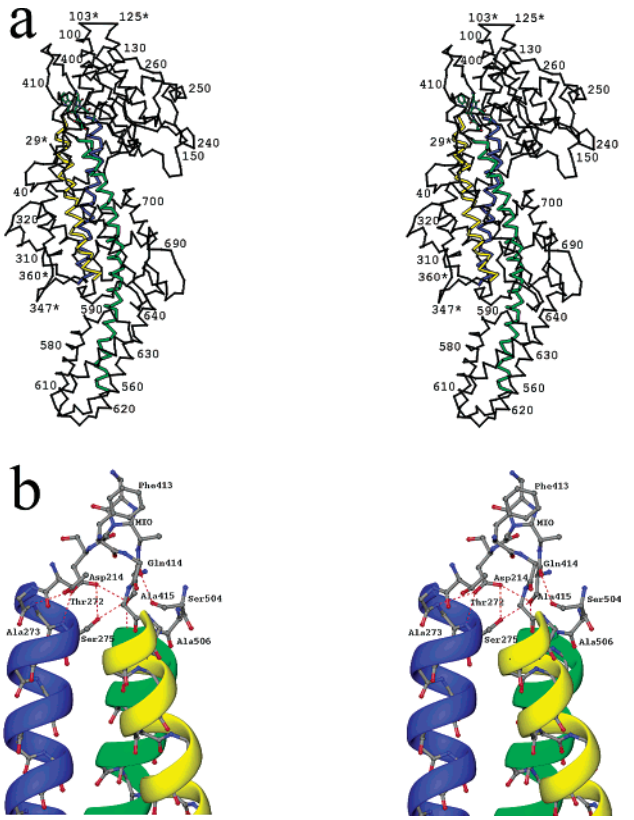


FIGURE 2: Tertiary structure of PAL. The three central core helices, leading to the active site, are colored blue, green, and yellow for emphasis. (a) Stereoview of the PAL monomer with residue numbering. Breaks in the polypeptide chain are indicated by asterisks. (b) Close-up stereoview of MIO and Phe413 interactions with the three central helices, polarized with their N termini directed toward the active site. Hydrogen bonds are indicated by dashed lines.

and PAL. Both PAL and HAL structures are based on a central core of nearly parallel α helices of varying lengths. The longest helix in PAL comprises 61 residues (505−565) and spans nearly the entire length of the monomer (Figure 2a). In PAL, there is only one section of β sheet longer than three residues (strands of residues 231−237 and 240−246); it resides in the funnel region leading to the active site. PAL differs from HAL by having an additional 215 residues, most of which are located in stretches at the N-terminal (57 residues) and C-terminal (98 residues) regions. These create a section of 155 residues extending above and below the main body of the structure in a "fan" arrangement (Figure 1a).

In higher plants, protease degradation, signaled by phosphorylation of PAL, is implicated as a regulatory mechanism of PAL activity (*42*). The phosphorylation site of the French bean PAL has been identified as Thr545, which does not have a corresponding Thr or Ser residue in the *R. toruloides* primary structure from an alignment, though Ser 562, Thr565, and Ser567 are in the vicinity. The latter residues reside near the "fan" region of PAL and are more than 40 Å away from the nearest MIO cofactor in the homotetramer, consistent with the lack of a direct effect of phosphorylation on PAL activity (*42*).

Within each monomer of PAL, three central α helices (residues 263−305, 415−445, and 505−565) form a triple-coiled coil similar to those observed in keratin and other fibrous proteins (*43*), suggesting a basis for a rigid central core (Figure 2). The three central helices are oriented with their resultant dipoles aligned similarly, creating an electropositive platform for cofactor MIO (Figure 2b). MIO (residues 211−213) is anchored to the three helices through noncovalent bonding: the Ser210 hydroxyl group shares hydrogen bonds with the amide NH of Ala506 of helix 505−565 and the OH group of Ser275 of the 263−305 helix; the carboxyl group of Asp214 forms a complex hydrogen-bonding network with the hydroxyl groups of Ser210 and Ser275 and the amide NH groups of Ser275 and Val274 of helix 263−305; and the pinnacle of the third helix (residues 415−445) connects to Phe413, whose ring system appears nearly ring-stacked with the ring of MIO and apparently functions to stabilize the putative NH_2 adduct (vide infra). Covalent linkages of MIO to the PAL backbone are to a loop on the cofactor's C terminus and to an α helix (residues 214−227) on its N terminus; the latter helix directs its positive pole toward the cofactor and the active site.

PAL Active Site. HAL and PAL contain the highly conserved Ala-Ser-Gly triad, which is converted autocatalytically, via internal cyclization and elimination of two water molecules, to the electrophilic prosthetic group 3,5-dihydro-5-methylidene-4*H*-imidazol-4-one (MIO). As in the HAL structure, the electron density well defines the MIO cofactor (residues 211−213 in PAL) that appears nearly planar (Figure 3). However, unlike the HAL structure, there is additional electron density connected to the methylidene group of the MIO cofactor in all 10 crystallographically independent subunits of the two PAL structures (Figure 3 and Supporting Information). The density is assigned as an NH_2 adduct to the MIO, because the penultimate purification step for PAL included hydrophobic interaction chromatography, where we estimate the protein eluted in the presence of 200 mM $(NH_4)_2SO_4$. The final step of purification (gel filtration), under our conditions, likely removed no more than 95% of the salt, leaving approximately 10 mM $(NH_4)_2SO_4$ in the crystallization drop. The inclusion of cinnamate in the crystallization drop would serve to stabilize the NH_2 adduct, owing to the determined order of release of product cinnamate before product NH_3 in the enzyme-catalyzed reaction (*44*). The phenyl ring of Phe413 is positioned to stabilize the putative NH_2 adduct of MIO in a nitrogen lone pair or NH to π-system acceptor interaction (*45*). Also, upon flipping the side chain amide of Asn270 by 180°, its OD1 becomes positioned to hydrogen-bond with the putative NH_2 adduct. There remains the possibility that the MIO adduct is something other than an NH_2 group (e.g., perhaps it is an OH group), which cannot be definitively determined from the limited resolution of the diffraction data. However, the HAL structure, which was determined from crystals grown in the absence of reaction products urocanate and ammonia and which has a similar structure to PAL, lacks evidence for an adduct to the MIO cofactor, suggesting that the adduct to the MIO of PAL is a consequence of crystallizing PAL in the presence of products cinnamate and ammonia. As well, the structure of HAL crystallized in the presence of inactivator L-cysteine shows that the methylidene group is covalently linked to the amino group of cysteine (*16*). Among the 10 crystallographically independent subunits of the two PAL structures, there is electron density to support the presence of cinnamate in only one subunit (Figure 3 and

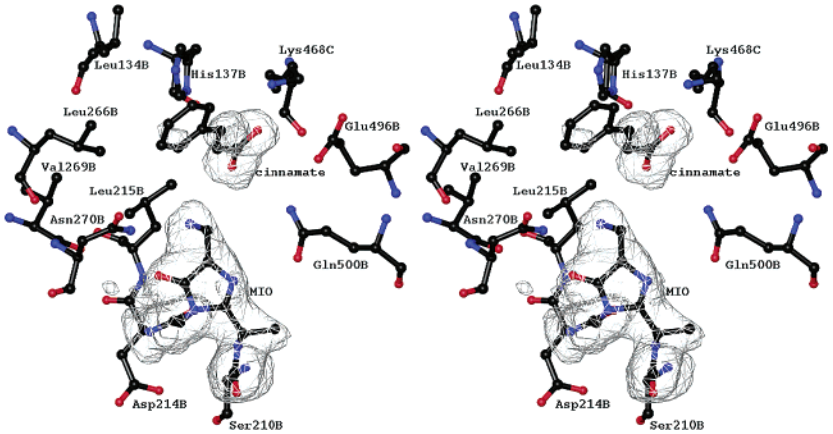


FIGURE 3: Stereoview of the PAL active site. Electron density ($2F_o - F_c$) corresponding to the MIO−NH_2 complex and product cinnamate in the active site of subunit B of the trigonal structure is shown. This is the only subunit (of the 10 subunits comprising the asymmetric units of the two structures reported) that has sufficient electron density for cinnamate, and it was refined as cinnamate in this subunit but not the others. All 10 subunits have sufficient electron density for the adduct connected to the methylidene group of MIO cofactor, refined as MIO−NH_2 in all 10 subunits. The $2F_o - F_c$ map is contoured at 1σ.

Supporting Information). The electron density covering part of cinnamate in one subunit and the lack of electron density for cinnamate in the other subunits may be due to multiple binding modes and/or high mobility of the product in the active site. Between pH 8.3 and 9.3, *trans*-cinnamate is a competitive inhibitor with respect to substrate phenylalanine with a K_i value of 24 μM (*20*). If *trans*-cinnamate has a similar binding constant under our crystallization conditions, the enzyme active sites would have been 98% occupied.

Although the MIO cofactor and its anchoring helices are all contained within a single monomer, a complete active site comprises residues from three separate monomers. Thus, the four monomers in the PAL tetramer create four active sites, each consisting of an MIO group sitting at the base of a funnel of residues leading from the active site to bulk solvent. In HAL the funnel is capped by two loops that enclose the active site, precluding access to bulk solvent, but in PAL the two loops (residues 104−124 and 348−359) are not resolved in the determination, suggesting that they are highly mobile. This difference between the HAL and PAL structures provides a likely mechanism for opening and closing the active site for substrate access.

In addition to the association of α helices with the MIO cofactor, most of the active-site residues of PAL are members of α helices, whose positive poles are directed toward the active site: in the active site with cofactor MIO formed by residues from subunit B, Leu215 belongs with the helix of residues 214−227 of subunit B; Leu266, Val269, and Asn270 belong with the helix of residues 263−305 of subunit B; Tyr363 and Arg366 belong with the helix of residues 367−393 of subunit D; Phe413 belongs with the helix of residues 415−445 of subunit B; Lys468, Gly469, and Ile472 belong with the helix of residues 467−486 of subunit C; and Glu496 and Gln500 belong with the helix of residues 505−565 of subunit B (Figure 4a). Direction of the positive poles of six α helices toward the active site of PAL suggests a potential means of supporting the development of a carbanion intermediate in the mechanism of Scheme 1, without invoking the development of a carbocation intermediate as in the mechanism of Scheme 2. A positively charged environment would disfavor the development of the carbocation of Scheme 2. A more detailed assessment of the

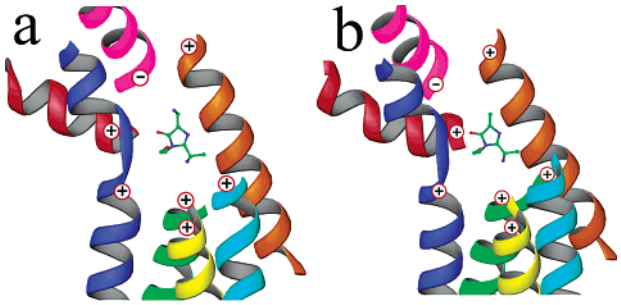


FIGURE 4: Seven α helices associated with the active sites of PAL and HAL. Shown are the six positive poles and the one negative pole of the seven α helices, directed toward active-site residues, with cofactor MIO shown for reference. (a) PAL. Clockwise from the negative pole of helix 126−139 of subunit B are the helices 467−486 of subunit C, 367−393 of subunit D, 415−445 of subunit B, 505−565 of subunit B, 263−305 of subunit B, and 214−227 of subunit B. (b) HAL. The respective order of helices is 71−84, 381−400, 284−309, 331−361, 419−453, 188−230, and 145−158.

mechanisms described by Schemes 1 and 2 is gleaned from models of substrate docked into the active site of PAL.

Models of L*-Phenylalanine in the Active Site of PAL.*[2] Differential crystal packing forces might be expected to impart structural differences. However, an overlay of active-site residues from the 10 crystallographically independent subunits of PAL shows that the positions of the residues, relative to that of cofactor MIO, are similar to one another (Figure 5); therefore, observations made on one active site should not differ much from those made on the others. For modeling purposes we chose the active site comprising the MIO formed in subunit B of the trigonal form, because it is from the higher-resolution structure and it contains electron density for a presumed cinnamate ion as shown in Figure 3.

With the protein kept as a rigid body, L-phenylalanine was placed into the active site of PAL to approximate the formation of compounds **1** and **4**, the first intermediates of Schemes 1 and 2, respectively, (Figure 6). The models remove the NH_2 adduct from the MIO cofactor of PAL, avoid steric conflicts between phenylalanine and the protein where possible, and place the substrate in a position close to the

[2] Coordinates of the models shown in Figures 6 and 8 are available from the corresponding authors.

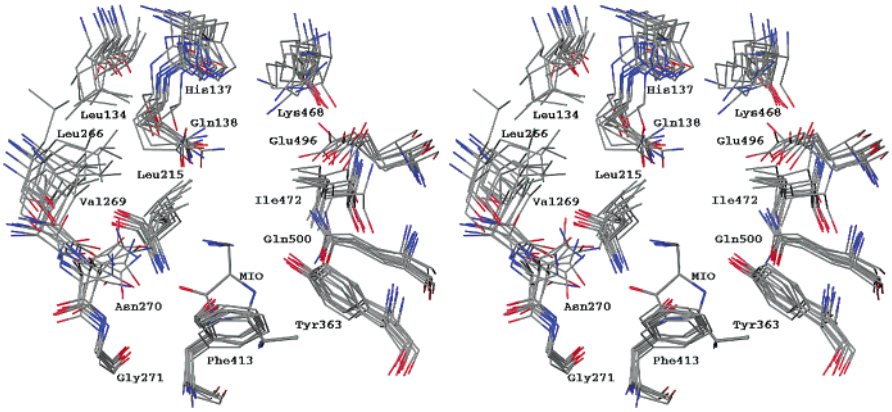


FIGURE 5: Overlay of ten PAL active sites. The stereoview includes the active sites belonging to the 10 crystallographically independent subunits of PAL contained in the two structures reported in this work. The active sites are centered on the MIO−NH_2 complex.

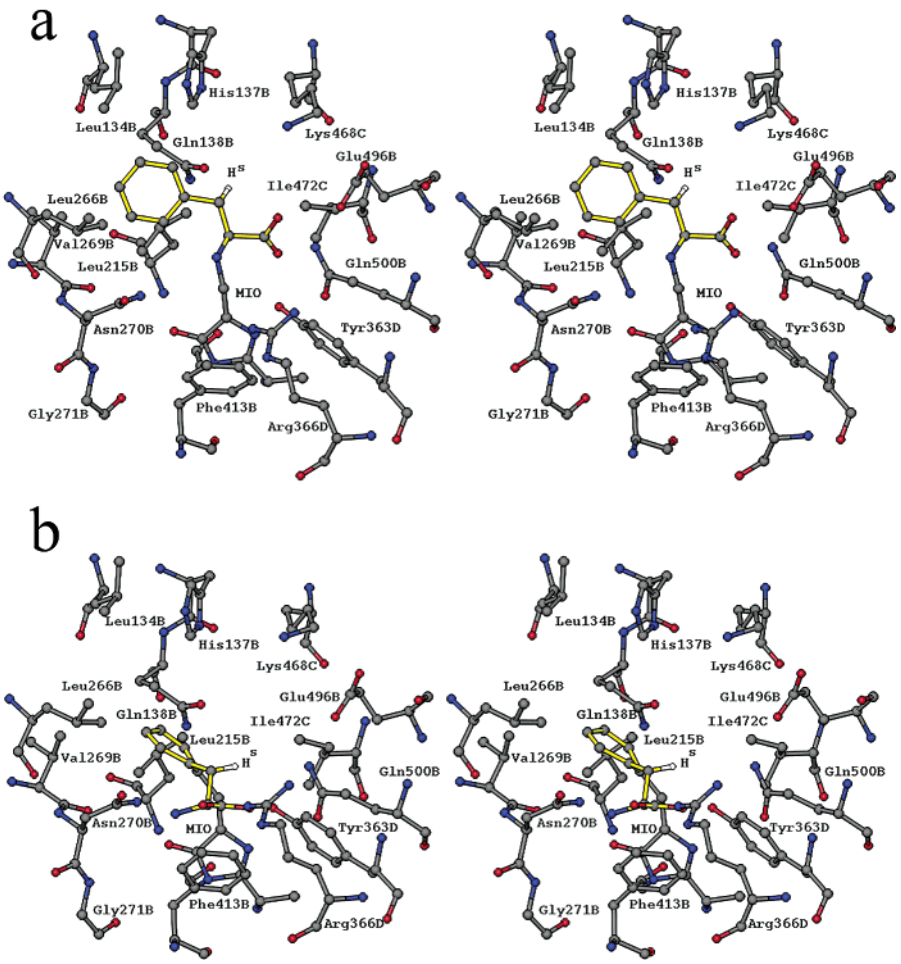


FIGURE 6: Stereoview models of substrate L-phenylalanine docked in the active site of PAL.[2] Subunit B of the trigonal structure was used for the docking. Bonds of substrate phenylalanine are colored yellow. (a) Docking of phenylalanine into the PAL active site to approximate intermediate **1** of Scheme 1. (b) Docking of phenylalanine into the PAL active site to approximate intermediate **4** of Scheme 2. The subunit origin of the residues is indicated by the suffix to the residue name (e.g., His137B is from subunit B).

methylidene group of MIO for forming a covalent bond with the cofactor (corresponding to that in intermediates **1** and **4**). Thus, the distance between the methylidene group of MIO and the amino group phenylalanine is 1.46 Å (ideal bond length = 1.46 Å) in Figure 6a, and the distance between the methylidene group and the C2 atom of the phenyl ring of phenylalanine is 1.53 Å (ideal bond length = 1.55 Å) in Figure 6b. The model of Figure 6b is based on the superposition of the carboxyl group of the substrate onto the location of an ordered SO_4^{2-} group found in the HAL crystal structure (*14*) so that the carboxyl group of substrate shares a salt bridge with an Arg side chain, similar to the models of histidine in the active site of HAL (*14*) and phenylalanine in the active site of the homology model of PAL (*26*). In the model that is in accord with Scheme 1 (Figure 6a), the SO_4^{2-} group is also considered an anion-friendly site, but instead of placing the carboxyl group of substrate there, this site is assumed to have a role in stabilizing the phenolate anion of Tyr363 formed after the residue donates a proton to the NH_2 leaving group of intermediate **3**; the carboxyl

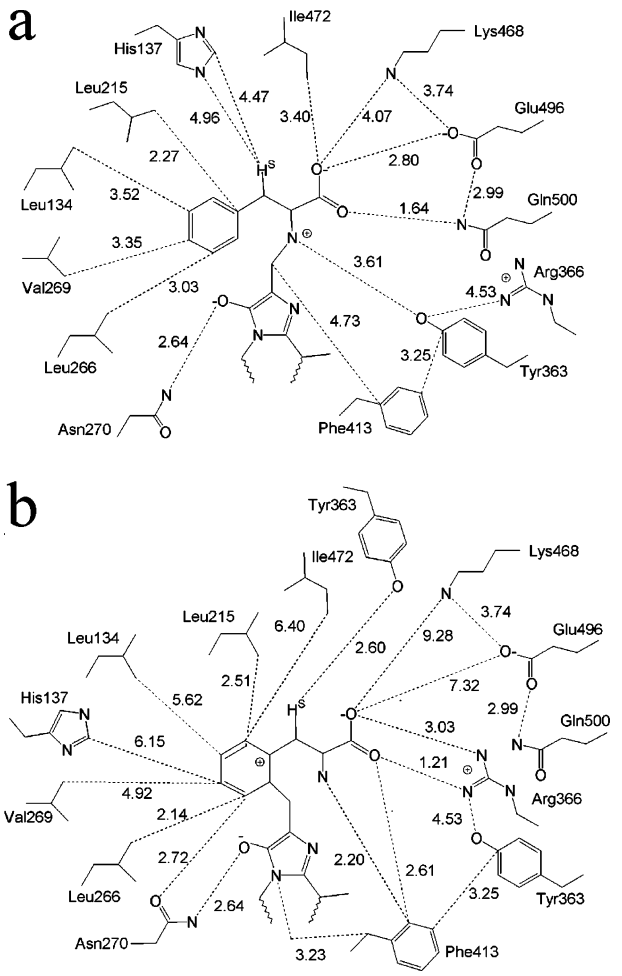


FIGURE 7: Schematics of interactions between substrate L-phenylalanine and residues in the active site of PAL. (a) Contacts between substrate L-phenylalanine and active site in the model of Figure 6a corresponding to the mechanism of Scheme 1. (b) Contacts between substrate L-phenylalanine and active site in the model of Figure 6b corresponding to the mechanism of Scheme 2. Closest contacts of residues with the docked substrate are indicated along with other contacts. Distances are in angstroms.

group of substrate is placed in an accommodating, hydrophilic pocket of the active site.

Inspection of the models in Figure 6 reveals that there are fewer steric conflicts between the docked substrate and the protein in panel a than in panel b: for example, in panel a, residues within 2.6 Å of substrate are Leu215 and Gln500; those in panel b are Leu215, Leu266, Phe413, and Arg366. A major feature that distinguishes the two models is that the side chains of the highly conserved lipophilic residues (Leu134, Leu215, Leu266, and Val269) form an accommodating pocket for the phenyl group of substrate in panel a, whereas a much less appealing pocket for the phenyl group is presented in the model of panel b. Specifically, the side chains of Leu134, Leu215, Leu266, and Val269 are 3.52, 2.27, 3.03, and 3.37 Å, respectively, from the phenyl group of substrate in Figure 6a, whereas in panel b the respective distances are 5.62, 2.51, 2.14, and 4.91 Å. Also, the carboxyl group of substrate appears much better accommodated by the model of panel a than that of panel b, by having interactions with three highly conserved, hydrophilic residues (Lys468, Glu496, and Gln500) rather than one (Arg366). Schematic drawings of Figure 6, indicating contact distances, are shown in Figure 7. Table 2 lists the active-site residues

Table 2: Conservation of Residues in the Active Sites of PAL and HAL and Effects of Single-Site Mutations on Catalysis[a]

residue		no. of exceptions (replacement residue)[b]		mutation to residue[b] (V_{maxwt}/V_{maxmut})	
PAL	HAL	PAL	HAL	PAL	HAL
Leu134	Leu79	8 (M)	41 (I)		
His137	His83	6 (R, N)	3 (L, S)		L (18 000)[c]
Ala211	Ala142	3 (S)	4 (C)		
Ser212	Ser143	0	0	A (5000)[d]	A (1160)[e]
Gly213	Gly144	0	0		
Leu215	Leu146	0	0		
Leu266	Leu191	0	3 (V)		
Val269	Leu194	6 (L)	69 (I, V)		
Asn270	Asn195	5 (M, T)	0	A (2600)[e]	A (24 300)[e]
Tyr363	Tyr280	5 (H)	1 (F)	F (235)[e]	F (55)[c]
Arg366	Arg283	5 (Q, V)	0	A (130)[f]	I (7840),[e] K (110)[e]
Phe413	Phe329	0	4 (G, A)	A (340)[e,f]	A (500)[e]
Lys468	Met382	0	1 (A)		
Gly469	Ile383	3 (A)	4 (V, G)		
Ile472	Val386	6 (V)	4 (Y, C)		
Glu496	Ser410	3 (S, D)	5 (C, N)		
Gln500	Glu414	0	2 (Q)	A (605),[c,e] E (6)[c,e]	A (101 250),[e] Q (34 700)[e]

[a] Residue numbering for PAL is from *R. toruloides*; 195 PAL sequences were compared in an alignment. Residue numbering for HAL is from *Pseudomonas putida*; 152 HAL sequences were compared in an alignment. [b] Single-letter designation of the residue change is indicated. [c] Data from ref *46*. [d] Data from ref *12*. [e] Data from ref *11*. [f] Data from ref *26*.

shown in Figure 6 and their corresponding residues in HAL from a sequence alignment. It also indicates most of the active-site residues of PAL and HAL are highly conserved or have conservative replacements according to separate multiple-sequence alignments and the effects of site-directed mutagenesis that have been reported for PAL and HAL (*11*, *12*, *26*, *46*). Most of the amino acid changes in the active site produce relatively large decreases in V_{max}, particularly considering that the change of the methylidene group of the MIO to a methyl group (Ser to Ala mutation) deteriorates V_{max} by factors of 5000 (PAL) and 1100 (HAL).

Within the context of Scheme 1, the mechanism of enzyme action can be postulated to proceed as follows. The funnel leading to the active site is rich in basic amino acids (His408, His409, His307, His312, His344, His345, Arg333, Arg359, Arg362, and Arg366), which could serve to favor binding of the NH_2 form of substrate phenylalanine over the NH_3^+ form. The NH_2 form of substrate is needed to form **1**, and this likely accounts for the p$K_a \sim 9$ observed for *V*/*K* (*20*) as suggested (*20*). The positively charged side chain of Lys468 (connected to helix 467−486 to augment its positive charge) could have a role in recognizing the carboxyl group of substrate (by forming a salt bridge) when it is located in the mouth of the funnel, and upon flipping the side chain downward it would chaperone the substrate to its reactive position for sharing additional interactions (hydrogen bonds) between the substrate's carboxyl group and the side chains of Glu496 and Gln500 (attached to the positive pole of helix 505−565) as shown in the model of Figure 6a. Lys468 is strictly conserved and the adjacent residue is almost always a Gly, which would increase the mobility of Lys468 in its role as chaperone for substrate. The chaperoning role of Lys468 would have the dual purpose of placing the substrate's NH_2 group near the MIO for forming **1** and ensuring

that the carboxylate group of substrate does not have an opportunity to react nonproductively with the methylidene of MIO by forming an ester. In these roles, the other pK_a ~ 9, determined for *V*/*K* (*20*), would be assigned to the positively charged form (active form) of Lys468. To promote formation of intermediate **1**, the side-chain NH_2 group of Asn270 of helix 263−305 is positioned to share a hydrogen bond with the developing enolate oxygen of MIO, increasing the electropositivity of the MIO methylidene group. Abstraction of the pro-*S* hydrogen from C3 of substrate to form **2** is catalyzed by His137, which is 5 Å away (ND1 to H^S) in the model and would need to move nearer for serving this function: His137 is part of α helix 126−139 that is connected to loop regions on both its C and N termini, which could allow movement. His137 is near the C terminus of the single α helix with its negative pole directed toward the active site; this would increase the basicity of His137 in its role as a general base. His137 would be active when its imidazole group is in the neutral form, consistent with the reported pK_a ~ 7, determined for V_{max} in the PAL reaction (*47*).[3] Development and stabilization of carbanion **2** are provided by the phenyl ring of the substrate whose electrons are polarized away from the carbanion site from interactions with Leu215 of helix 214−227 and Leu266 and Val269 of helix 253−305 (both helix dipoles direct their positive poles toward the residues), with some assistance from the electron-withdrawing capability of the MIO group (that is enhanced by being at the apex of the positive poles of three α helices) and electron withdrawal from the substrate's carboxyl group by residues associated with the N termini of α helices. The electron-withdrawing capability of the MIO group would assist the breakage of the C−N bond of substrate in proceeding from **2** to **3** and product, *trans*-cinnamate. Once the C−N bond is broken, product cinnamate is free to leave the active site with its carboxyl group escorted by Lys468, which is consistent with the kinetic results that indicate cinnamate is the first product to leave the enzyme (*44*). The NH_2 adduct of **3** is stabilized by sharing a hydrogen bond with OD1 of Asn270 and by an interaction with the phenyl ring of Phe413, which is near the positive pole of helix 415−445, suggesting a nitrogen lone pair to π system interaction. Breakdown of **3** to NH_3 and MIO is promoted by donation of a proton from the hydroxyl group of Tyr363 to the NH_2 group of **3**. Tyr363 is near the N terminus of helix 367−393, whose positive pole would increase the acidity of Tyr363 in its role as a general acid. The resting state of Tyr363 would be active when its hydroxyl group is in the neutral form, consistent with the pK_a ~ 10, determined for V_{max} in the PAL reaction (*47*).[3] The phenolate form of Tyr363 would be stabilized by the nearby guanidinium group of Arg366 (made more positive by being near the positive pole of helix 367−393). Exit of NH_3 and cinnamate from the active site would be facilitated by opening the loops that cover the active-site funnel leading to bulk solvent; exposure to bulk solvent would return the enzyme to its resting state. The residues shown in Figure 6a are either strictly conserved or highly conserved among species of PAL (Table 2). Of the residues in Table 2 and Figure 6a, which are within the volume of the active site selected for presentation, all have been accounted for in the mechanism of Scheme 1, with the exception of Ile472.

The mechanism of Scheme 2 is difficult to reconcile with the pH studies of the PAL-catalyzed reaction and the determined pK_a values; for example, the pK_a ~ 9 for *V*/*K*, assigned to the amine of substrate in the mechanism of Scheme 1. In the mechanism of Scheme 2, the carboxyl group of substrate phenylalanine forms a salt bridge with Arg366 once it enters the active site, as shown by the model of Figure 6b. Formation of intermediate **4** (Figure 6b) is promoted by the electrophilic character of MIO that is enhanced by the polarized α helices with which it is associated; it is opposed by the positive poles of helices 214−227 and 263−305 through the proximity of Leu215 and Leu266 to the phenyl ring of substrate. The enolate of **4** is stabilized by hydrogen-bond sharing with the nearby amide of Asn270. In proceeding from **4** to **5**, the phenol group of Tyr363, activated by Arg366, abstracts the nearby pro-*S* hydrogen from C3 of substrate (2.6 Å from O to H^S). The third and final step has Asn270 reclaiming full possession of its proton that leads to the collapse of **5** to products. Thus, the model of Figure 6b accounts for roles of the highly conserved residues Asn270, Tyr363 and Arg366 but not the highly conserved residues His137, Glu496, Gln500, and Lys468, which are >6 Å away from substrate phenylalanine in the model. Clearly, the substrate could be oriented differently than that shown in Figure 6b, which is purposely similar to that reported for a homology model of PAL (*26*). However, all attempts at improvement have yielded models of **4** in the PAL active site that are inferior to the model of **1** in the PAL active site (Figure 6a) from a structural point of view.

Models of L*-Histidine in the Active Site of HAL.*[2] PAL and HAL active sites are similar and share many of the same residue types (Table 2). In this work we point out that, similar to PAL, the MIO of HAL sits atop a platform of three α helices with their positive poles directed toward the active site. Also, we point out that active-site residues in HAL that correspond to those in PAL (Table 2) belong to six α helices whose positive poles are directed toward the active site with one helix directing its negative pole there, as found in PAL (Figure 4b). Previous models of substrate histidine in the HAL active site have assumed the mechanism of Scheme 2 (*14*, *16*). The modeling results on PAL, which favor the mechanism of Scheme 1 over that of Scheme 2, prompted us to explore models of HAL considering both mechanisms. As with the PAL models, HAL was kept as a rigid body, waters and the SO_4^{2-} group were removed, and L-histidine was placed into the active site of HAL to approximate the formation of compounds **1** and **4** (replacing the phenyl ring of phenylalanine with the imidazole ring of histidine), corresponding to the first intermediates of Schemes 1 and 2, respectively (Figure 8). The model of Figure 8b is similar to that reported (*14*, *16*); it places the carboxyl group in the position occupied by the SO_4^{-2} anion. In the model of Figure 8a, the site of the SO_4^{2-} group is considered an anion-friendly site, being near the hydroxyl group of Tyr363, which becomes a tyrosyl anion in the reaction coordinate of Scheme 1. The schematic diagrams of Figure 9 indicate that there are few steric conflicts in the models of Figure 8: for example, there are two interactions closer than 2.0 Å in both

[3] Value determined for PAL from potato tubers. This enzyme does not display Michaelis−Menten kinetics, and pK_a values were estimated from a plot of V_{satn} (velocity saturated with substrate at each pH) versus pH (*47*).

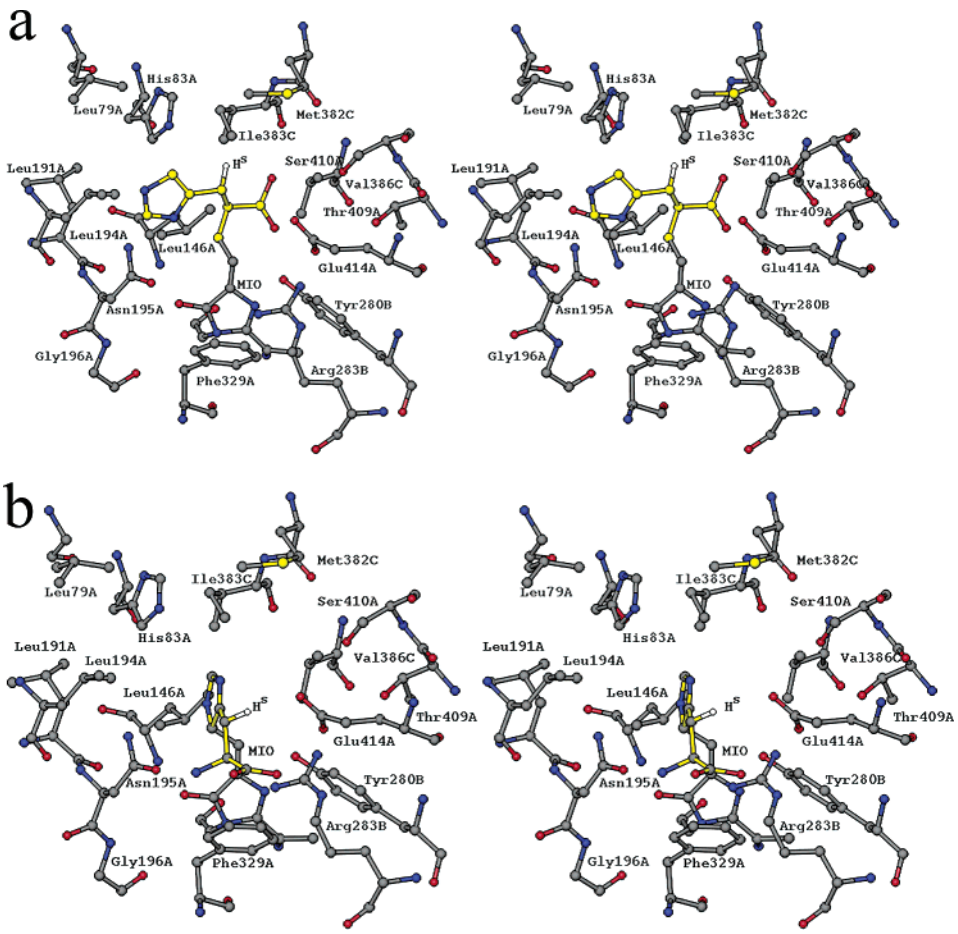


FIGURE 8: Stereoview models of substrate L-histidine docked in the active site of HAL.[2] Bonds of substrate histidine are colored yellow. (a) Docking of histidine into the HAL active site to approximate intermediate **1** of Scheme 1. (b) Docking of histidine into the HAL active site to approximate intermediate **4** of Scheme 2. The subunit origin of the residues is indicated by the suffix to the residue name.

models. Possible mechanisms of the HAL-catalyzed reaction according to Schemes 1 and 2 are similar to those described for PAL.

HAL has a highly conserved Met382 (of helix 381−400) that corresponds to Lys468 in PAL, but the proposed roles would be similar according to the model of Figure 8a and Scheme 1; the side chain of Met382 (near the positive pole of helix 381−400) would escort the carboxylate group of substrate L-histidine (chelated to a divalent metal) to its productive position. Once positioned, divalent metal would chelate the carboxyl group of substrate, the sulfur group of Met382, the carboxyl group of Glu414, and the hydroxyl group of Ser410 that form an appealing site for the cation. The amino group of substrate would add to the methylidene group of cofactor MIO (assisted by Asn195 sharing a hydrogen bond with the developing enolate anion of MIO) to form the intermediate corresponding to **1**. Removal of the pro-*S* hydrogen from C3 of substrate to form the intermediate corresponding to **2** is catalyzed by His83, which is 3.3 Å away (ND1 to H^S) in the model. His83 is near the C terminus of the single α helix (helix 71−84) with its negative pole directed toward the active site, and this would support His83 in its role as a general base. Development and stabilization of the carbanion corresponding to **2** is by the imidazole ring of the substrate whose electrons are polarized away from the carbanion site from interactions with Leu146 of helix 145−158 and Leu191 and Leu194 of helix 188−230 (both helices direct their positive poles toward the residues) with some assistance from the electron-withdrawing capability of the MIO group (enhanced by being at the apex of the positive poles of three α helices) and electron withdrawal from the carboxyl group of substrate by its associated divalent metal and residues belonging to the N termini of helices. The electron-withdrawing capability of the MIO group would assist breakage of the C−N bond of substrate in proceeding from **2** to **3** and product *trans*-urocanate. The NH_2 adduct of the intermediate corresponding to **3** is stabilized by sharing a hydrogen bond with the OD1 of Asn195 and by an interaction with the phenyl ring of Phe329 (near the positive pole of helix 331−361). Breakdown of **3** to NH_3 and MIO is promoted by donation of a proton from the hydroxyl group of Tyr280 (near the positive pole of helix 284−309) to the NH_2 group of **3**. The phenolate of Tyr363 would be stabilized by the nearby guanidinium group of Arg283 of the same helix. Exit of NH_3 and urocanate from the active site would be assisted by opening of the loops that cover the active-site funnel leading to bulk solvent.

In the mechanism of Scheme 2, the carboxyl group of substrate L-histidine forms a salt bridge with Arg366 once it enters the active site, as shown in the model of Figure 8b. Formation of intermediate **4** (Figure 6b) is promoted by the electrophilic character of MIO that is enhanced by the polarized α helices to which it is attached; it is opposed by the positive pole of helix 145−158 through the proximity of Leu146 to the imidazole ring of substrate. The enolate of **4** is stabilized by hydrogen-bond sharing with the nearby

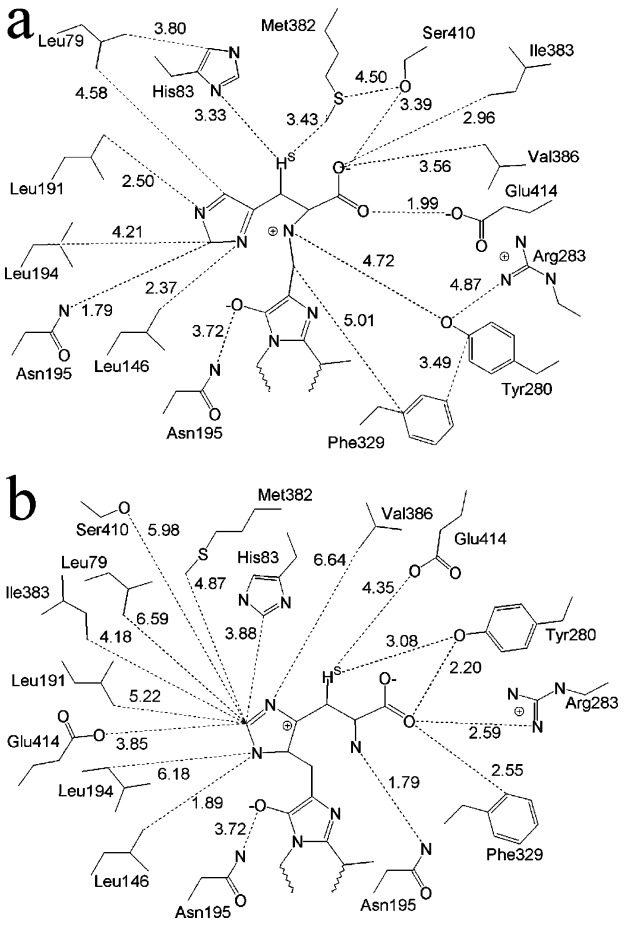


FIGURE 9: Schematics of interactions between substrate L-histidine and residues in the active site of HAL. (a) Contacts between substrate L-histidine and active site residues in the model of Figure 8a corresponding to the mechanism of Scheme 1. (b) Contacts between substrate L-histidine and active site residues in the model of Figure 8b corresponding to the mechanism of Scheme 2. Closest contacts of residues with the docked substrate are indicated along with other contacts. Distances are in angstroms.

amide of Asn270. In proceeding from **4** to **5**, the phenol group of Tyr280 (attached to helix 284−309), activated by Arg283, removes the nearby pro-*S* hydrogen from C3 of substrate (3.1 Å from O to H^S). The final step has Asn270 reclaiming possession of its proton that leads to the collapse of **5** to products. As in the case of the PAL models of Figure 6, the model of HAL according to the mechanism of Scheme 1 (Figure 8a) accounts for the roles of more of the conserved active-site residues than that of the mechanism of Scheme 2 (Figure 8b).

Conclusions. The existence of a carbanion intermediate, in the reaction coordinate of PAL acting on dihydrophenylalanine (*20*), is undisputed, and an E1cb-like mechanism is explicit in both mechanisms of Schemes 1 and 2. In the E1cb-like mechanism, PAL faces a major challenge in adjusting pK_a values to support the development of a carbanion intermediate. Off the enzyme, the pK_a of the C3 methylene group of phenylalanine is high and would need to drop ∼30 pH units to match the pK_a of an unactivated enzyme general base such as histidine or tyrosine for proton abstraction from C3 to proceed. The mechanism of Scheme 2, which invokes the development of a carbocation intermediate on the phenyl ring of phenylalanine at the expense of breaking ring aromaticity, was devised, in part, to account for the ability of the catalyst to stabilize a carbanion intermediate when there was not sufficient means known of doing so. Its appeal is that addition of the phenyl ring of phenylalanine to the methylidene group of MIO to form a carbocation provides a mechanism to lower the pK_a of C3 considerably so that it would match that of an enzyme base. However, substitution of Ala for Ser of the Ala-Ser-Gly triad to change the methylidene group of cofactor to a methyl group, so that the carbocation intermediate (**4** of Scheme 2) cannot be formed, erodes V_{max} by only a factor of 10^3−10^4 (Table 2). Borohydride treatment of wild-type PAL or HAL, to reduce the methylidene group of MIO to a methyl group for achieving the same effect, causes similar erosion in V_{max} (*12*, *48*). Therefore, the methylidene group of MIO accounts for up to 4 but leaves ∼26 pK_a units unaccounted for in the discrepancy between the C3 of substrate and the enzyme base. This assessment of relevant experimental data raises serious doubts as to the existence of carbocation **4** in the enzyme-catalyzed reaction. Clearly, the MIO cofactor has roles in accelerating catalysis, but fully equalizing the pK_as is not its responsibility.

A viable means of equalizing the pK_a values of the C3 methylene group of phenylalanine and the enzyme base to favor the development of carbanion intermediate **2** of Scheme 1 is through the influence of dipole moments of α helices. The MIO cofactor of PAL is atop three α helices that direct their positive poles toward the environment of the active site and strengthen the MIO electrophile, and most active-site residues are associated with six α helices that direct their positive poles toward the active site for supporting specific roles in catalysis. Electron-withdrawing capability associated with active-site residues, which are modeled to interact with the phenyl ring of substrate, would polarize electron density away from the ring system to lower the pK_a of the C3 group. MIO and other active-site residues associated with the positive poles of dipoles could assist in the pK_a challenge. The basicity of the enzyme general base (His137) is strengthened by being near the negative pole of a helix. There is no experimental accounting for the specific influences of the individual residues and their associated helices in adjusting the pK_a values of C3 and enzyme general base for equalization, but it can be evaluated by dividing the ∼30 units of pK_a by the seven α helices associated with the active site to arrive at an average contribution of ∼4 pK_a units per helix dipole. The latter value is in line with the drop of ∼5 pK_a units of an active-site cysteine (pK_a of 3.4) that has been attributed to an individual helix dipole interaction in thiol−disulfide oxidoreductase (*49*).

The positive poles of α helices and their associated residues in the active site of PAL appear to specifically disfavor the development of carbocation **4** in the mechanism of Scheme 2 and the proposed roles of active-site residues within the mechanism. If the mechanism of Scheme 2 were to prevail, it would appear to constitute the sole example of protein architecture working against an enzyme catalyzing its own physiological reaction. In addition to the dipoles of α helices in the PAL structures, our modeling results of substrate phenylalanine in the PAL structure favor the mechanism of Scheme 1 over that of Scheme 2: the model of intermediate **1** is well accommodated by the conserved hydrophobic and hydrophilic regions of the active site (Figure 6a), and conserved active-site residues, with pK_a values

similar to those determined for the enzyme-catalyzed reaction, are well positioned for participating in catalysis. In common with structures of other enzymes that catalyze E1cb-like mechanisms with active-site residues poised to remove both the proton and the leaving group from modeled substrate to approximate the appearance of an E2 mechanism (*50*), PAL has His137 and the MIO group in place for achieving the *anti* elimination of the elements of ammonia (Figure 6a). Whereas the model of Figure 6a is imperfect in displaying a near attack conformation (NAC) (*51*), it is close. Taken together, the PAL structure appears fully compatible with the mechanism of Scheme 1 (model of Figure 6a), and it appears incompatible with the mechanism of Scheme 2 (model of Figure 6b).

In this work we point out that, in common with PAL, HAL has a similarly disposed bundle of three α helices that direct their positive poles toward MIO and a similar arrangement of six α helices associated with active-site residues that direct their positive poles toward the active site and one helix that has its negative pole directed toward His83 to strengthen the general base. The HAL active site well accommodates the intermediate that corresponds to **1** of Scheme 1 (Figure 8a), and many of the same residue types, as implicated in the PAL reaction, are well positioned to serve similar catalytic roles in the HAL reaction. The HAL model of Figure 8a displays a more appealing view of a NAC with respect to the eliminated elements of ammonia than the PAL model (Figure 6a). Our analysis of PAL, favoring the mechanism of Scheme 1 over Scheme 2, applies to HAL. The multiple α helices associated with PAL and HAL active sites cannot seriously be considered biological accidents, because the enzymes catalyze mainstream, physiological reactions subject to selection pressures for catalytic efficiency.

An adduct to the MIO functionality was found in all 10 of the active sites refined from the two PAL crystals grown in the presence of *trans*-cinnamate and NH_4 ions. The data were not collected at high enough resolution to positively assign the identity as NH_2, but it seems the most likely one. If the MIO adduct in PAL is genuinely that of NH_2, the structure provides a view of the PAL active site occupied by intermediate **3** and cinnamate in accordance with the mechanism of Scheme 1. The NH_2 adduct, directed toward the phenyl ring of Phe413 and the side-chain amide of Asn270, provides a mechanism for separating the reaction products for promoting the removal of cinnamate from the active site before it has a chance to re-form phenylalanine in a back reaction with intermediate **3**. Together, seven helix dipoles in the PAL active site appear to serve multiple duties in catalysis: increasing the electrophilicity of the MIO cofactor, increasing the positive charge of Lys468 to strengthen its interaction with the carboxylate group of substrate, polarizing the phenyl ring of substrate to lower the pK_a of C3 and stabilize the carbanion intermediate, increasing negative charge on His137 to strengthen the general base, activating Phe413 to attract the NH_2 adduct and separate it from cinnamate, and withdrawing charge from Tyr363 to strengthen the general acid.

Dipole moments originating from α helices have long been implicated in roles serving enzyme catalysis; most notably, in presenting binding sites for ions (mainly anions) and in modulating pK_a values of active-site residues, cofactors, and substrates (*49*, *52*−*54*). α Helices are abundant in PAL and HAL. Of the 24 α helices longer than three residues in PAL, seven are directed toward the active site and implicated in catalysis; in HAL, of the 17 α helices longer than three residues, seven are implicated in catalysis and their residues align with the residues in the seven α helices of PAL. The multiple helix dipoles implicated in the PAL and HAL reaction mechanisms are testimony to the difficulty of activating the nonacidic C3 of substrate for proton abstraction and orchestrating the remaining reaction events. Polarizing the aromatic ring of substrate through interactions with active-site residues connected to the N termini of α helices appears to be the means by which PAL and HAL are best equipped to activate C3 of substrate. Complementarities of the enzyme active-site residues and the aromatic ring of substrate would be crucial for realizing the effect of the helix dipoles, and the effectiveness of the helix dipoles would have governance over substrate specificity, particularly through an E1cb mechanism with its demanding pK_a requirement.

Recently, another enzyme has been discovered to contain cofactor MIO (*55*, *56*). Tyrosine aminomutase (TAM) catalyzes the transfer of the amino group from C2 to C3 of substrate L-tyrosine in producing β-tyrosine, a reaction that has been shown to proceed through a *p*-hydroxycinnamate intermediate and to have significant tyrosine ammonia lyase activity. Further structure−function studies should shed light on how TAM suppresses the ammonia lyase function to favor the aminomutase function; elucidate features in PAL, HAL, and TAL that govern substrate specificity; and define the role of the metal cofactor in HAL (and possibly TAL). Together, such studies should serve to better define the roles of the MIO cofactor and its associated helix dipoles in these fascinating catalysts that act on amino acid substrates.

ACKNOWLEDGMENT

We thank Drs. Katharine Gibson, Der-Ing Liao, Mark Nelson, Will Marshall, Xiao-Song Tang, Jean-Francois Tomb, and Ya-Jun Zheng (all of DuPont) for helpful discussions; Dr. R. M. Sweet for the beam time at NSLS; and Dr. Jeff Bell (Accelyris, Inc.) for help with the MIO topology and parameter library patches for the refinement with CNX.

SUPPORTING INFORMATION AVAILABLE

Figures showing $F_o - F_c$ electron density maps for the MIO−NH_2 complex and cinnamate in the PAL active site. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hanson, K. R., and Havir, E. A. (1981) Phenylalanine Ammonia-Lyase, in *The Biochemistry of Plants* (Conn, E., Ed.) Vol. 7, pp 577−625, Academic Press, New York.
2. Rosler, J., Krekel, F., Amrhein, N., and Schmid, J. (1997) Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity, *Plant Physiol. 113*, 175−179.
3. 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.
4. Gatenby, A. A., Sariaslani, S., Tang, X. S., Qi, W. W., and Vannelli, T. (2002) Bioproduction of para-hydroxycinnamic acid, U.S. Patent 6,368,837 B1.

5. Turner, A. M., Simpson, A., Mcinnes, R., and Howell, P. L. (1997) Human argininosuccinate lyase: a structural basis for intragenic complementation, *Proc. Natl. Acad. Sci. U.S.A. 94*, 9063−9068.
6. Shi, W., Dunbar, J., Jayasekera, M. M. K., Viola, R. E., and Farber, G. K. (1997) The structure of L-aspartate ammonia-lyase from *Escherichia coli*, *Biochemistry 36*, 9136−9144.
7. Jayasekera, M. M. K., Shi, W., Farber, G. K., and Viola, R. E. (1997) Evaluation of functionally important amino acids in L-aspartate ammonia-lyase from *Escherichia coli*, *Biochemistry 36*, 9145−9150.
8. Levy, C. W., Buckley, P. A., Sedelnikova, S., Kato, Y., Asano, Y., Rice, D. W., and Baker, P. J. (2002) Insights into enzyme evolution revealed by the structure of methylaspartate ammonia lyase, *Structure 10*, 101−113.
9. Peterkofsky, A. (1962) The mechanism of action of histidase: amino-enzyme formation and partial reactions, *J. Biol. Chem. 237*, 787−795.
10. Rettig, M., Sigrist, A., and Rétey, J. (2000) Mimicking the reaction of phenylalanine ammonia-lyase by a synthetic model, *Helv. Chim. Acta 83*, 2246−2265.
11. Poppe, L. (2001) Methylidene-imidazolone: a novel electrophile for substrate activation, *Curr. Opin. Chem. Biol. 5*, 512−524.
12. Schuster, B., and Rétey, J. (1995) The mechanism of action of phenylalanine ammonia-lyase: the role of prosthetic dehydroalanine, *Proc. Natl. Acad. Sci. U.S.A. 92*, 8433−8437.
13. Skolaut, A., and Rétey, J. (2001) 1,4-Dihydro-L-phenylalanine—its synthesis and behavior in the phenylalanine ammonia-lyase reaction, *Arch. Biochem. Biophys. 393*, 187−191.
14. Schwede, T. F., Rétey, 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.
15. Baedeker, M., and Schulz, G. E. (2002) Autocatalytic peptide cyclization during chain folding of histidine ammonia-lyase, *Structure 10*, 61−67.
16. Baedeker, M., and Schulz, G. E. (2002) Structures of two histidine ammonia-lyase modifications and implications for the catalytic mechanism, *Eur. J. Biochem. 269*, 1790−1797.
17. Rétey, J. (2003) Discovery and role of methylidene imidazolone, a highly electrophilic prosthetic group, *Biochim. Biophys. Acta 1647*, 179−184.
18. Langer, B., Langer, M, and Rétey, J. (2001) Methylidene-imidazolone (MIO) from histidine and phenylalanine ammonia-lyase, *Adv. Protein Chem. 58*, 175−214.
19. Rétey, J. (1996) Enzymatic catalysis by Friedel−Crafts-type reactions, *Naturwissenschaften 83*, 439−447.
20. Hermes, J. D., Weiss, P. M., and Cleland, W. W. (1985) Use of nitrogen-15 and deuterium isotope effects to determine the chemical mechanism of phenylalanine ammonia-lyase, *Biochemistry 24*, 2959−2967.
21. Gloge, A., Langer, B., Poppe, L., and Rétey, J. (1998) The behavior of substrate analogues and secondary deuterium isotope effects in the phenylalanine ammonia-lyase reaction, *Arch. Biochem. Biophys. 359*, 1−7.
22. Furuta, T., Takahashi, H., and Kasuya, Y (1990) Evidence for a carbanion intermediate in the elimination of ammonia from L-histidine catalyzed by histidine ammonia-lyase, *J. Am. Chem. Soc. 112*, 3633−3636.
23. Lewandowicz, A., Jemielity, J., Kanska, M., Zon, J., and Paneth, P. (1999) Tritium secondary kinetic isotope effect on phenylalanine ammonia-lyase-catalyzed reaction, *Arch. Biochem. Biophys. 370*, 216−221.
24. Lewis, B. E., and Schramm, V. L. (2003) Binding equilibrium isotope effects for glucose at the catalytic domain of human brain hexokinase, *J. Am. Chem. Soc. 125*, 4785−4798.
25. Xing., S., and Cekan, S. Z. (1983) Isotope effects of tritium in a ligand−antibody binding, *Int. J. Appl. Radiat. Isot. 34*, 957−958.
26. Röther, D., Poppe, L., Morlock, G., Viergutz, S., and Rétey, J. (2002) An active site homology model of phenylalanine ammonia-lyase from *Petroselinum crispum*, *Eur. J. Biochem. 269*, 3065−3075.
27. Hendrickson, W. A., Pahler, A., Smith, J. L., Satow, Y., Merritt, E. A., and Phizackerley, R. P. (1989) Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation, *Proc. Natl. Acad. Sci. U.S.A. 86*, 2190−2194.
28. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
29. Sheldrick, G. M. (1998) Location of Heavy Atoms by Automated Patterson Interpretation, in *Direct Methods for Solving Macromolecular Structures* (Fortier, S., Ed.) pp 131−141, Kluwer Academic Publishers, Dordrecht, The Netherlands.
30. Weeks, C. M., and Miller, R. (1999) The design and implementation of SnB version 2.0, *J. Appl. Crystallogr. 32*, 120−124.
31. De La Fortelle, E., and Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods, *Methods Enzymol. 276*, 472−494.
32. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement, *J. Appl. Crystallogr. 30*, 1022−1025.
33. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
34. McRee, D. E. (1999) XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density, *J. Struct. Biol. 125*, 156−165.
35. Brünger, A. T., and Warren, G. L. (1998) Crystallography and NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
36. Brünger, A. T. (1992) The free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures, *Nature 355*, 472−474.
37. Word, J. M., Lovell, S. C., Richardson, J. S., and Richardson, D. C. (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation, *J. Mol. Biol. 285*, 1735−1747.
38. Lu, G. (2000) TOP: a new method for protein structure comparisons and similarity searches, *J. Appl. Crystallogr. 33*, 176−183.
39. Bailey, S. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
40. Carson, M. (1997) Ribbons, *Methods Enzymol. 277*, 493−505.
41. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
42. Allwood, E. G., Davies, D. R., Gerrish, C., Ellis, B. E., and Bolwell, G. P. (1999) Phosphorylation of phenylalanine ammonia-lyase: evidence for a novel protein kinase and identification of the phosphorylated residue, *FEBS Lett. 457*, 47−52.
43. Mitraki, A., Miller, S., and van Raaij, M. J. (2002) Review: conformation and folding of novel beta-structural elements in viral fiber proteins: the triple beta-spiral and triple beta-helix, *J. Struct. Biol. A137*, 236−247.
44. Havir, E. A., and Hanson, K. R. (1975) L-Phenylalanine ammonia-lyase (maize, potato, and *Rhodotorula glutinis*). Studies of the prosthetic group with nitromethane, *Biochemistry 14*, 1620−1626.
45. Sarkhel, S., Rich, A., and Egli, M. (2003) Water-nucleobase "stacking": H−π and lone pair−π interactions in the atomic resolution crystal structure of an RNA pseudoknot, *J. Am. Chem. Soc. 125*, 8998−8999.
46. Röther, D., Poppe, L., Viergutz, S., Langer, B., and Rétey, J. (2001) Characterization of the active site of histidine ammonia-lyase from *Pseudomonas putida*, *Eur. J. Biochem. 268*, 6011−6019.
47. Havir, E. A., and Hanson, K. R. (1968) L-Phenylalanine ammonia lyase. II. Mechanism and kinetic properties of the enzyme from potato tubers, *Biochemistry 7*, 1904−1914.
48. Langer, M., Pauling, A., and Rétey, J. (1995) The role of dehydroalanine in catalysis by histidine ammonia lyase, *Angew. Chem., Int. Ed. Engl. 34*, 1464−1465.
49. Phillips, B., and Glockshuber, R. (2002) Randomization of the entire active-site helix α1 of the thiol-disulphide oxidoreductase DsbA from *Escherichia coli*, *J. Biol. Chem. 277*, 43050−43957.
50. Jordan, D. B., Zheng, Y.-J., Locket, B. A., and Basarab, G. S. (2000) Stereochemistry of the enolization of scytalone by scytalone dehydratase, *Biochemistry 39*, 2276−2282.
51. Bruice, T. C., and Lightstone, F. C. (1999) Ground state and transition state contributions to the rate of intramolecular and enzymic reactions, *Acc. Chem. Res. 32*, 127−136.
52. Hol, W. G. (1985) The role of the alpha-helix dipole in protein function and structure, *Prog. Biophys. Mol. Biol. 45*, 149−195.
53. Thoden, J. B., Holden, H. M., Zhuang, Z., and Dunnaway-Mariano, D. (2002) X-ray crystallographic analyses of inhibitor and substrate complexes of wild-type and mutant 4-hydroxybenzoyl-CoA thioesterase, *J. Biol. Chem. 277*, 27468−27476.

54. Trickey, P., Wagner, M. A., Jorns, M. S., and Mathews, F. S. (1999) Monomeric sarcosine oxidase: structure of a covalently flavinylated amine oxidizing enzyme, *Structure 7*, 331−345.
55. Christenson, S. D., Liu, W., Toney, M. D., and Shen, B. (2003) A novel 4-methylideneimidazole-5-one-containing tyrosine aminomutase in enediyne antitumor antibiotic C-1027 biosynthesis, *J. Am. Chem. Soc. 125*, 6062−6063.
56. Christenson, S. D., Wu, W., Spies, M. A., and Shen, B. (2003) Kinetic analysis of the 4-methylideneimidazole-5-one-containing tyrosine aminomutase in enediyne antitumor antibiotic C-1027 biosynthesis, *Biochemistry 42*, 12708−12718.

BI049053+