# Crystallographic Studies of Phosphonate-Based α-Reaction Transition-State Analogues Complexed to Tryptophan Synthase[†,‡]

Aristidis Sachpatzidis, Chris Dealwis, Jodi B. Lubetsky, Po-Huang Liang, Karen S. Anderson, and Elias Lolis*

*Department of Pharmacology, Yale University School of Medicine, New Haven, Connecticut 06520*

*Received April 2, 1999; Revised Manuscript Received July 23, 1999*

ABSTRACT: In an effort to use a structure-based approach for the design of new herbicides, the crystal structures of complexes of tryptophan synthase with a series of phosphonate enzyme inhibitors were determined at 2.3 Å or higher resolution. These inhibitors were designed to mimic the transition state formed during the α-reaction of the enzyme and, as expected, have affinities much greater than that of the natural substrate indole-3-glycerol phosphate or its nonhydrolyzable analogue indole propanol phosphate (IPP). These inhibitors are ortho-substituted arylthioalkylphosphonate derivatives that have an $sp^3$-hybridized sulfur atom, designed to mimic the putative tetrahedral transition state at the C3 atom of the indole, and lack the C2 atom to allow for higher conformational flexibility. Overall, the inhibitors bind in a fashion similar to that of IPP. Glu-49 and Phe-212 are the two active site residues whose conformation changes upon inhibitor binding. A very short hydrogen bond between a phosphonate oxygen and the Ser-235 hydroxyl oxygen may be responsible for stabilization of the enzyme−inhibitor complexes. Implications for the mechanism of catalysis as well as directions for more potent inhibitors are discussed.

De novo biosynthesis of aromatic amino acids is a unique feature of prokaryotic and lower eukaryotic organisms. The genes encoding proteins in these biosynthetic pathways are present in bacteria, fungi, and plants, but have been deleted from animals during the course of evolution. Therefore, enzymes that function in these pathways could be molecular targets for the development of new antibacterial, antifungal, or selective herbicidal agents with limited animal toxicity. Two inhibitors of the shikimate pathway are among the examples that support the belief that targeting these pathways could be fruitful. (6*R*)-6-Fluoro-5-enolpyruvylshikimate-3-phosphate is an antibacterial and antifungal inhibitor of chorismate synthase (*1*), and the herbicidal agent glyphosate is an inhibitor of 5-enolpyruvylshikimate-3-phosphate synthase (*2*).

Tryptophan synthase (TRPS)[1] (EC 4.2.1.20) catalyzes the final two reactions in the de novo biosynthesis of L-tryptophan (reviewed in refs *3*−*5*). The bacterial protein is a bifunctional enzyme with an $\alpha_2\beta_2$ subunit stoichiometry. The α-subunit catalyzes a retroaldol reaction in which indole 3-glycerol phosphate (IGP) is cleaved to indole and D-glyceraldehyde 3-phosphate (G3P) (α-reaction). The nascent indole reacts with the activated L-serine in a pyridoxal phosphate (PLP)-dependent reaction in the active site of the β-subunit to form L-tryptophan and water (β-reaction). The three-dimensional structure of tryptophan synthase (TRPS) from *Salmonella typhimurium* has been determined by X-ray crystallography and reveals that a 25 Å long hydrophobic tunnel connects the active sites of the α- and β-subunits (*6*). This hydrophobic tunnel is believed to have two roles. It allows indole, the product of the α-reaction and the substrate of the β-reaction, to diffuse through the protein, thereby preventing its loss by diffusion through the cell membrane. The tunnel also is involved in the allosteric interactions that synchronize the reactions in the two subunits.

In addition to the unliganded enzyme, structures of mutants and wild-type enzymes complexed to a variety of substrates and inhibitors have been reported (*6*−*11*). These crystallographic studies have been complemented by extensive kinetic analyses of wild type and site-directed mutants of the enzyme in providing a detailed view of the enzymatic mechanism of the reaction. In particular, the formation of the aminoacrylate at the β-subunit active site activates the cleavage of indole glycerol 3-phosphate at the α-site with the subsequent release of indole into the tunnel (*12*−*15*).

The wealth of structural, kinetic, and mechanistic information that is available for tryptophan synthase makes it an attractive target for inhibitor design. In an effort to develop herbicidal agents, a series of phosphonate inhibitors of tryptophan synthase were designed, synthesized, and tested for inhibition of enzyme activity in an in vitro assay and for herbicidal activity in a biological assay (*63*).[2] The design of these compounds was based on the putative transition state of the natural substrate IGP formed during the α-reaction of

[†] A.S. was supported by fellowships from the A. S. Onassis Foundation and the United States Educational Foundation in Greece. J.B.L. was supported by NIH Training Grant GM07223. K.S.A. was supported by NIH Grant AI 44630, and E.L. was supported by NIH Grant AI 43838.

[‡] The coordinates have been deposited at the Brookhaven Data Bank under file names 1C29, 1C8V, 1CX9, 1C9D, and 1CW2.

* To whom correspondence should be addressed. Telephone: (203) 785-6233. Fax: (203) 785-7670. E-mail: elias.lolis@yale.edu.

[1] Abbreviations: AA, aminoacrylate; G3P, D-glyceraldehyde 3-phosphate; GP, α-glycerol 3-phosphate; IGP, indole 3-glycerol phosphate; IPP, indole propanol phosphate; FIPP, 5-fluoroindole propanol phosphate; $NAD^+$, oxidized nicotinamide adenine dinucleotide; PLP, pyridoxal 5′-phosphate; TRPS, tryptophan synthase; rmsd, root-mean-square deviation; ⟨*esd*⟩, mean coordinate error; LBHB, low-barrier hydrogen bond; NMR, nuclear magnetic resonance; TSA, transition-state analogue.

10.1021/bi9907734 CCC: $18.00 
Published on Web 09/09/1999

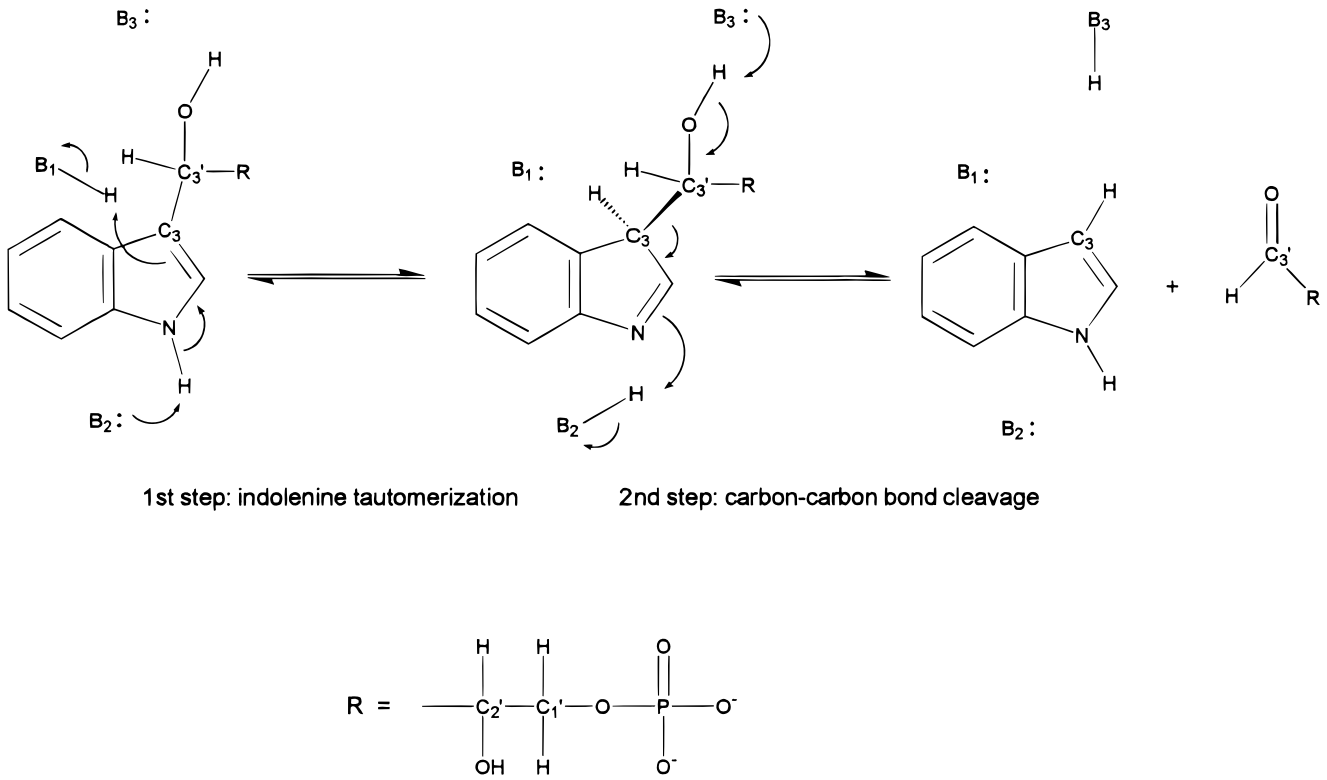


FIGURE 1: Mechanism of the α-reaction catalyzed by tryptophan synthase. In the first step, the base $B_2$ polarizes the N−H bond of indole and abstracts the proton. The acid $B_1H$ protonates the C3 atom and facilitates the electronic rearrangement that leads to the formation of the indolenine tautomer. In the second step, the actual C−C bond cleavage is catalyzed by the base $B_3$, which abstracts the proton from the C1′ hydroxyl of the glyceryl side chain of indole 3-glycerol phosphate. This event leads to the formation of indole and D-glyceraldehyde 3-phosphate.

the enzyme (*63*). According to the proposed mechanism (Figure 1) (*3*, *4*), the rate of the α-reaction is enhanced by "push−pull" general acid−base catalysis (*16*). Specifically, the cleavage of the C3−C3′ carbon−carbon bond in IGP is activated by tautomerization of the indole ring to yield the indolenine tautomer. Two groups, $B_1H$ and $B_2$, facilitate this tautomerization. Asp-60 has been identified as $B_2$ (*6*, *8*, *11*, *17*, *18*), but the identity of $B_1H$ remains unclear. Glu-49 could function as $B_1H$, but the available evidence is inconclusive (*4*). $B_1H$ protonates the indole ring at C3, while $B_2$ abstracts the proton on N1 of the ring to yield the tautomeric form. The actual aldol cleavage step is catalyzed by $B_3$, identified as Glu-49 (*9*), which abstracts a proton from the C3′ hydroxyl of the glyceryl side chain, an event that leads to the formation of indole and G3P.

Here we describe structural studies on five arylthioalkylphosphonate transition-state analogues designed to inhibit the α-reaction (*63*).[2] To establish the molecular basis of inhibition by these agents, the crystal structures of the corresponding complexes have been determined at 2.3 Å or higher resolution. The information provided here has implications for the mechanism of catalysis and can be used as a guide for the design of more potent inhibitors.

## MATERIALS AND METHODS

*Chemicals.* The following tryptophan synthase inhibitors were provided by American Cyanamid: 4-(2-hydroxyphenylthio)-1-butenylphosphonic acid, 1:2 salt with isopropylamine (**1**); 4-(2-hydroxyphenylthio)butylphosphonic acid, 1:2 salt with diisopropylamine (**2**); 4-(2-aminophenylthio)butylphosphonic acid (**3**); 4-(2-hydroxy-5-fluorophenylthio)butylphosphonic acid, 1:1 salt with diisopropylamine (**4**); and 4-(2-hydroxyphenylsulfinyl)butylphosphonic acid (**5**). The chemical synthesis of these compounds and a more complete characterization of their inhibitory activities has been published elsewhere (*63*).

*Enzyme Activity Measurements.* The expression and purification of the tryptophan synthase $\alpha_2\beta_2$ complex from *S. typhimurium* have been described previously (*19*). The ability of the arylthioalkylphosphonate compounds to inhibit the α-reaction of tryptophan synthase was confirmed by measuring the decrease of the initial rate at various concentrations of the inhibitors. The assay conditions that were used were those described by Creighton and Yanofsky (*20*) with some of the modifications introduced by Kirshner et al. (*21*). The inhibitors were tested at concentrations ranging from 0.1 to 5 $\mu$M. On the basis of these measurements, $IC_{50}$ values were calculated by using a nonlinear regression analysis of the direct data implemented in Kaleidagraph (Synergy Software, Reading, PA).

*Crystallization and X-ray Data Collection.* The protein−inhibitor complexes were prepared by mixing the individual components so that the final protein and inhibitor concentrations were 5−10 mg/mL and 1 mM, respectively. Crystals of the complexes were grown under conditions (50 mM Bicine, 1 mM NaEDTA, 0.8−1.5 mM spermine, and 12% PEG 4000 adjusted to pH 7.8 with NaOH) that were modified from the original protocol to crystallize the unliganded

[2] Langevine, C. M., and Finn, J. M. (1997) Arylthioalkyl- and arylthioalkenylphosphonic acids and derivatives thereof useful as herbicidal agents, U.S. Patent 5,635,449.

Table 1: Crystal Parameters and Data and Refinement Statistics

| | TRPS−**1** | TRPS−**2** | TRPS−**3** | TRPS−**4** | TRPS−**5** |
|---|---|---|---|---|---|
| crystal parameters | | | | | |
| unit cell dimensions *a*, *b*, *c* (Å) | 182.3, 58.9, 67.4 | 183.8, 60.8, 68.2 | 182.7, 59.3, 67.3 | 184.2, 60.5, 67.8 | 183.3, 59.9, 67.7 |
| unit cell dimension $\beta$ (deg) | 94.4 | 94.4 | 94.5 | 94.4 | 94.5 |
| data statistics | | | | | |
| resolution (Å) | 42.2−2.3 | 45.8−2.2 | 42.7−2.3 | 45.9−2.3 | 45.7−2.0 |
| no. of collected reflections | 245222 | 223028 | 110360 | 95281 | 238914 |
| no. of unique reflections | 29830 | 35625 | 30288 | 31780 | 47432 |
| completeness (total/high) (%) | 92.6/85.2 | 93.5/87.2 | 90.6/70.8 | 95.2/79.6 | 95.8/91.8 |
| $R_m$ (total/high) (%)[a] | 7.6/16.1 | 7.8/18.3 | 11.7/27.6 | 7.3/17.3 | 5.4/13.4 |
| $\langle I/\sigma(I)\rangle$ (total/high)[a] | 12.9/3.9 | 7.4/2.9 | 8.0/4.1 | 11.2/4.5 | 15.4/5.9 |
| refinement statistics | | | | | |
| resolution range (Å) | 30−2.3 | 30−2.2 | 30−2.3 | 30−2.3 | 30−2.0 |
| no. of reflections with $F > 2\sigma(F)$ | 29402 | 35371 | 29619 | 31553 | 43982 |
| no. of protein atoms[b] | 4928 | 4928 | 4928 | 4928 | 4928 |
| no. of waters | 162 | 160 | 157 | 188 | 184 |
| no. of other atoms[c] | 32 | 32 | 32 | 33 | 33 |
| $R_{free}$, $R_{work}$ (%) | 27.5/20.6 | 26.9/20.8 | 27.6/21.0 | 27.1/22.1 | 25.3/21.7 |
| rmsd for bonds/angles (Å/deg) | 0.010/1.74 | 0.010/1.81 | 0.008/1.68 | 0.009/1.69 | 0.011/1.86 |
| disallowed ($\Phi$ and $\Psi$) residues[d] | Phe-212 | | | Phe-212 | Phe-212 |
| $\langle B\rangle$ (mc/sc/wat) (Å$^2$)[e] | 14.6/16.7/15.7 | 11.6/15.1/18.8 | 16.3/19.6/17.1 | 13.9/17.4/15.3 | 15.9/19.6/22.9 |
| $\langle esd\rangle$ (Å)[f] | 0.25 | 0.22 | 0.28 | 0.25 | 0.23 |

[a] Completeness, $R_m$, and $\langle I/\sigma(I)\rangle$ are given for all data and for data in the highest-resolution shell. $R_m = \Sigma|I - \langle I\rangle|/\Sigma I$. [b] No unambiguous electron density was found for the following residues in the atomic model: $\alpha$1, $\alpha$190, $\alpha$191, $\alpha$268, $\beta$1, and $\beta$390−$\beta$397. [c] Other atoms represent PLP and a sodium ion in all cases and the corresponding phosphonate inhibitor in each complex. [d] Disallowed ($\Phi$ and $\Psi$) residues are defined as those residues with backbone torsion angles that are outside the allowed regions of the Ramanchandran diagram. [e] Mean thermal *B* factors are given for main chain (mc) and side chain (sc) protein atoms and water molecules (wat). [f] $\langle esd\rangle$ is the mean coordinate error estimated by the SIGMAA method.

enzyme (*6*). The crystals exhibit symmetry of space group *C*2 with an $\alpha\beta$ pair in the asymmetric unit.

Diffraction data were collected at a low temperature (140 K) on an Raxis IIC image plate system with CuK$\alpha$ X-rays generated from a Rigaku RU-200 rotating anode operating at 50 kV and 100 mA and equipped with a Yale double-mirror system. The crystal to detector distance was 100 mm and the oscillation range 1°. Reflection intensities were integrated with DENZO (*22*) and scaled with the CCP4 (*23*) suite of programs.

*Refinement.* The starting model for all five structures was the coordinate set of a refined model of native TRPS (PDB file name 1a5s) (*11*) without the cofactor PLP. X-PLOR, version 3.851 (*24*), was employed for all calculations. The graphics program O (*25*) was used for the display of electron density maps ($2F_{obs} - F_{calc}$ and $F_{obs} - F_{calc}$ difference syntheses at varying contour levels) and manual rebuilding of atomic models. The $R_{free}$ (*26*) was implemented from the beginning and its value used as a criterion for model improvement during the course of the refinement. After an initial round of rigid-body refinement, the model was subjected to a simulated annealing protocol starting at 4000 K. At this point, atomic models of the phosphonate inhibitor for each complex and of the common cofactor PLP that were generated and geometrically minimized with InsightII (Molecular Simulations Inc., San Diego, CA) were built into the corresponding electron density. Several rounds of slow cooling protocols with varying weights and starting temperatures, grouped and individual *B* factor refinement, and manual rebuilding followed. Placement of water molecules was achieved by selecting the peaks in $F_{obs} - F_{calc}$ difference maps that had heights of $>4\sigma$ and fulfilled hydrogen bonding criteria. A two-parameter bulk solvent correction (*27*) was applied, and this allowed low-resolution (5−30 Å) reflections to be used in the refinement. In the final stages of refinement, the coordinates and *B* factors of the atomic model were refined by using the conjugate gradient minimization algorithm. Data and refinement statistics are listed in Table 1.

## RESULTS AND DISCUSSION

*Rationale for Inhibitor Design.* The design strategy was based on the inhibitor indole propanol phosphate from the crystal structure of the TRPS−IPP complex and the presumed mechanism of the reaction. The first compounds prepared and tested for enzyme inhibitory and herbicidal activity were the phosphonate isosteres (*63*). The phosphonate group was chosen to provide a combination of desirable pharmacodynamic and pharmacokinetic properties. This group is expected to possess hydrogen-bonding capability similar to that of the phosphate group, and it is not susceptible to metabolic inactivation by phosphatases. The initial phosphonate analogues possessed enzyme inhibitory and herbicidal activity with micromolar $IC_{50}$s (*63*). To improve the potency of the phosphonate inhibitors, compounds were designed to mimic the transition state. The ortho-substituted arylthioalkylphosphonates are believed to share physicochemical (electronic and stereochemical) properties with the putative transition state of the $\alpha$-reaction and are likely to bind the enzyme with higher affinity than the unmodified substrate or product. During the chemical transformation in the $\alpha$-reaction, the C3 atom of indole undergoes a hybridization change from $sp^2$ in IGP to $sp^3$ in the transition state. An $sp^3$-hybridized sulfur atom replaces this tetrahedral carbon in all of the phosphonate inhibitors used in our study. In addition, the C2 atom of indole has been removed to allow for higher conformational flexibility. The $IC_{50}$s of the five compounds for inhibition of the $\alpha$-reaction of TRPS are provided in Figure 2.

*Quality of the Structures and Error Estimates.* The five structures reported in this study have been refined to 2.3 Å

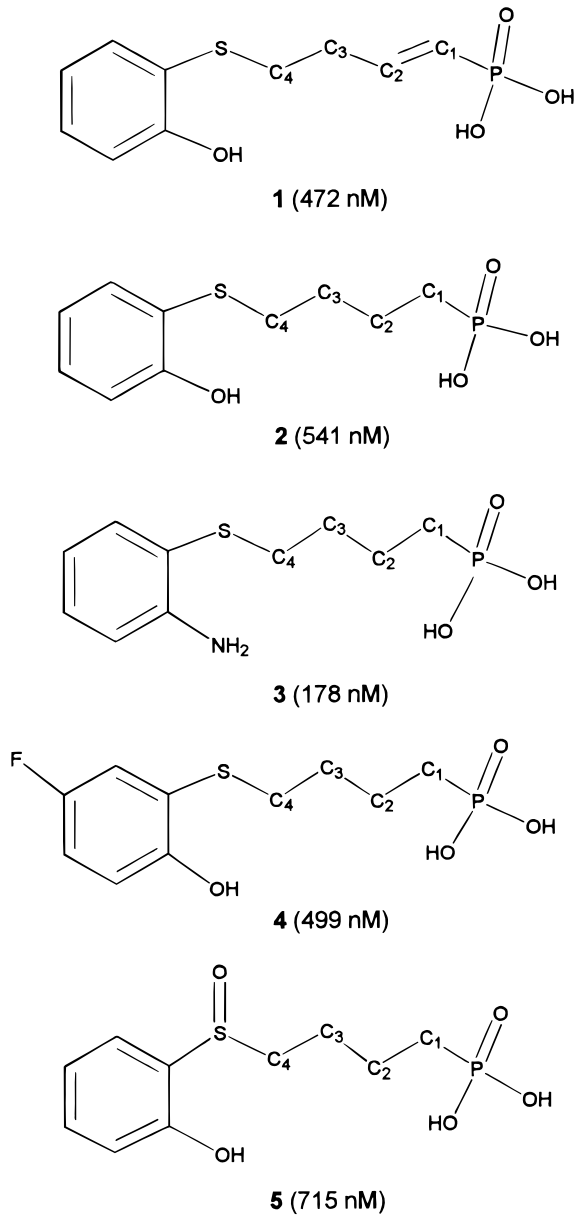


FIGURE 2: Chemical structures and $IC_{50}$ values of the transition-state analogue-based phosphonate inhibitors of tryptophan synthase. The $IC_{50}$ values are shown in parentheses next to the number representing each inhibitor: **1**, 4-(2-hydroxyphenylthio)-1-butenylphosphonic acid; **2**, 4-(2-hydroxyphenylthio)butylphosphonic acid; **3**, 4-(2-aminophenylthio)butylphosphonic acid; **4**, 4-(2-hydroxy-5-fluorophenylthio)butylphosphonic acid; and **5**, 4-(2-hydroxyphenylsulfinyl)butylphosphonic acid.

or higher resolution with $R_{free}$ values ranging from 25.3 to 28.1% and very good stereochemistry (Table 1). The protein residues that have been included in the final atomic models are α2−α189, α192−α267, and $\beta$3−$\beta$389. Electron density for the N- and C-terminal residues is not visible. Electron density exists between residues α189 and α192, but attempts to build the two intervening residues as a continuous peptide chain were unsuccessful. All residues, with the exception of αPhe-212, appear in the allowed regions of the Ramachandran plot. This residue is involved in inhibitor binding and is discussed in more detail below. As previously described, the α-subunit of TRPS is an eight-fold α/$\beta$ barrel, and the $\beta$-subunit is a two-domain protein with a unique fold (*6*).

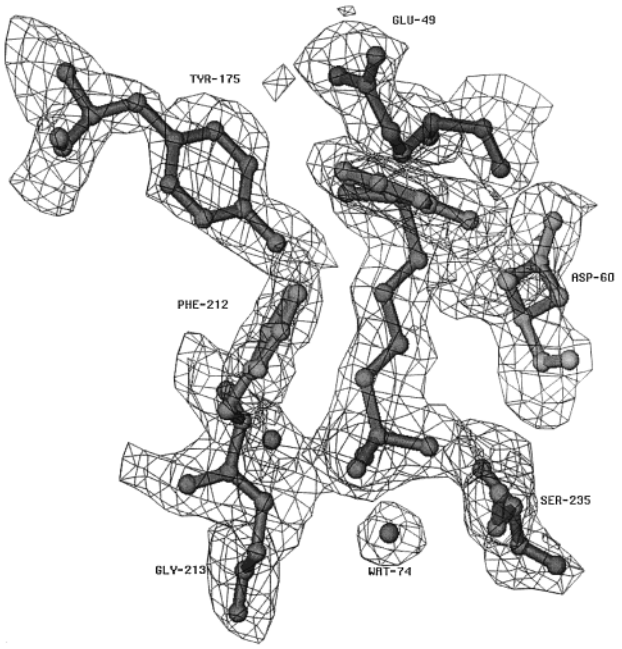


FIGURE 3: $2F_{obs} - F_{calc}$ $\sigma_A$-weighted simulated annealing omit electron density map (*29*, *31*) for the tryptophan synthase−inhibitor **5** complex at the α-subunit active site contoured at 1.0$\sigma$. A total of 265 protein and inhibitor atoms were omitted from the model before a final simulated annealing run and map calculation.

An estimate of the average coordinate error of the refined atomic models can be obtained either by the Luzzatti method (*28*) or the $\sigma_A$ plot (*29*, *30*). The $\sigma_A$-based estimates, which use normalized structure factors and are thus insensitive to scaling errors, are provided in Table 1. The Luzzatti plot gives comparable estimates for the structures. These error estimates are consistent with the rms differences obtained by pairwise least-squares superposition of the main chain atoms for the protein residues of the five inhibitor complexes. The rms differences range between 0.2 and 0.4 Å. Therefore, 0.4 Å can be considered an upper limit for the estimate of coordinate error in our structures, since the only difference is the nature of the inhibitor in the active site. The overall estimates of errors are useful for assessing the significance of comparisons between different structures. For assessing the accuracy of the positions of individual atoms in a structure, inspection of the electron density map is appropriate. Figure 3 presents the $\sigma_A$-weighted simulated annealing omit map for a representative inhibitor (**5**) and active site residues. The electron density indicates that the position of the atoms of each phosphonate inhibitor in the active site of TRPS is well-defined.

*Enzyme−Inhibitor Interactions.* As expected, the phosphonate inhibitors bind to the α-reaction active site. Potential hydrogen bonding interactions and relative distances from active site residues for the different inhibitors are shown in Figure 4A−E. Some interactions are common to all inhibitors, while others are unique. The phenyl ring and side chain (thiobutyl, thiobutenyl, or sulfinylbutyl groups) of all inhibitors make contact with a number of hydrophobic residues, including Phe-22, Leu-100, Leu-127, Phe-212, Leu-232, and the methyl group of Thr-183. This is very similar to the packing of the indole and propyl moieties of the substrate analogue FIPP (*11*). The alkylphosphonate portion of the inhibitors extends approximately at a right angle with the

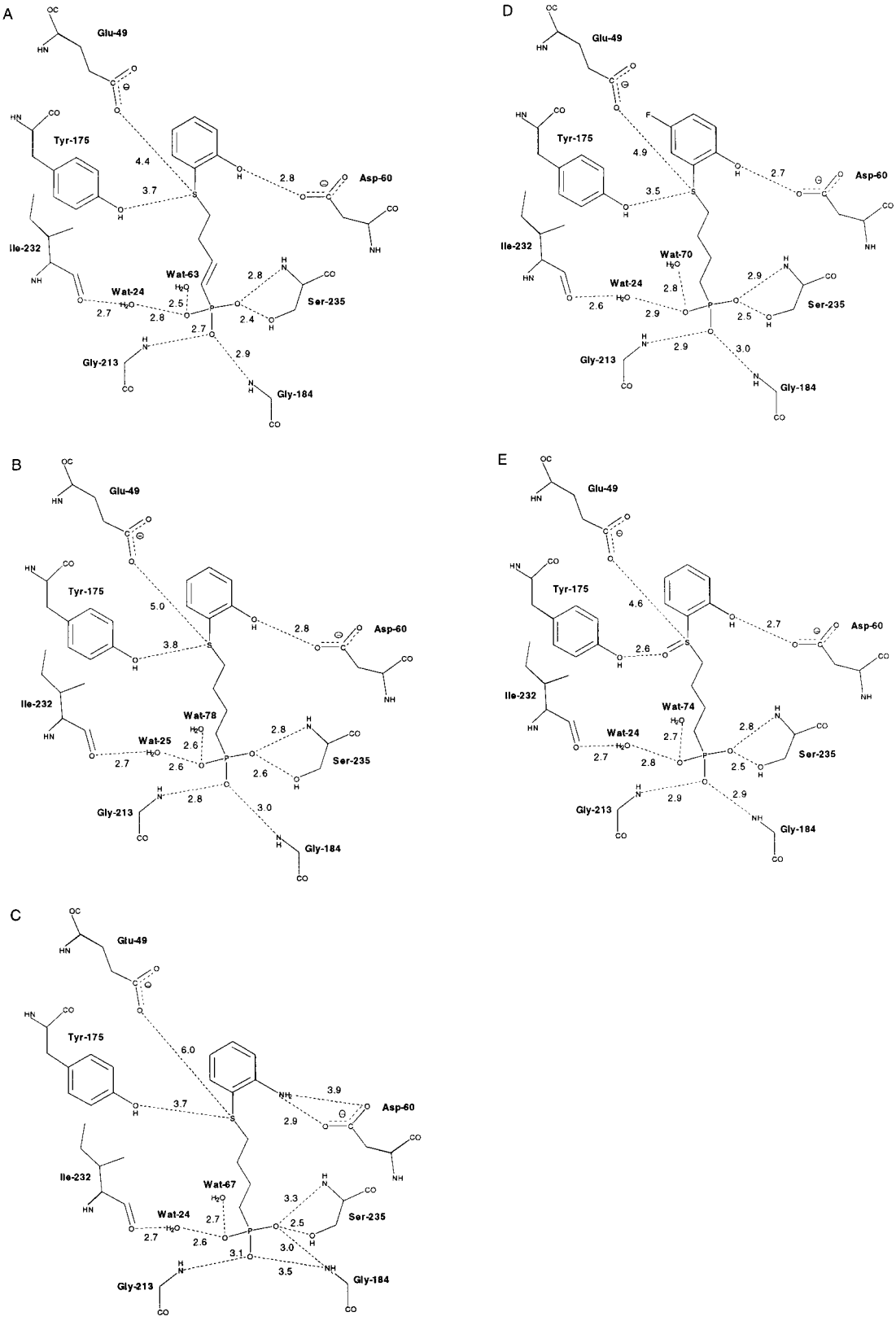


FIGURE 4: Schematic drawings of hydrogen bonding interactions between the five phosphonate inhibitors and putative catalytic residues at the α-subunit active site. Distances from Glu-49 and Tyr-175 to the sulfur atom of four of the inhibitors are also shown: (A) inhibitor **1** [4-(2-hydroxyphenylthio)-1-butenylphosphonic acid], (B) inhibitor **2** [4-(2-hydroxyphenylthio)butylphosphonic acid], (C) inhibitor **3** [4-(2-aminophenylthio)butylphosphonic acid], (D) inhibitor **4** [4-(2-hydroxy-5-fluorophenylthio)butylphosphonic acid], and (E) inhibitor **5** [4-(2-hydroxyphenylsulfinyl)butylphosphonic acid].

phenyl ring, and the phosphonate oxygens form hydrogen bonds with the hydroxyl group of Ser-235, two water molecules, and the main chain nitrogens of Gly-184, Gly-213, Gly-234, and Ser-235. The interaction with the hydroxyl of Ser-235 appears to be particularly strong in the complexes of TRPS with **1** and **3**−**5**. The significance of this interaction

is discussed below. As with IPP (and FIPP), no basic amino acid side chain makes direct interactions with the phosphonate. The side chain of Arg-179 is in the vicinity of the phosphonate group but at distances that are too large (4−5 Å) for any hydrogen bonding interaction to occur. However, in the low-dielectric environment within the protein, Arg-179 may be involved in substantial electrostatic interactions with the phosphonate. The phosphonate group also binds near the N-terminus of helix H8′ (terminology of ref *6*). The α-helix macrodipole may play a role in fine-tuning its fit to the active site as has been previously proposed (*6*). The ortho substituent of the phenyl ring consistently interacts with the carboxylate of putative catalytic residue Asp-60 (X−O distances range from 2.7 to 2.9 Å, where X is O or N).

In **1**, the rigidity introduced by the double bond does not perturb the potential for hydrophobic and van der Waals interactions, yet presumably favors binding due to entropic effects (i.e., fewer degrees of freedom are lost upon binding than in the case of a saturated C−C bond). Furthermore, in the rigid conformation imposed by the double bond, one of the phosphonate oxygens is brought into very close contact with the hydroxyl of Ser-235, forming a strong, possibly low-barrier, hydrogen bond (O···O interatomic distance refined to 2.4 Å) (*32*−*35*). The length of this hydrogen bond is the shortest among the inhibitors in this study and presumably reflects the strength of the interaction. The *o*-hydroxyl group of **2** forms a strong interaction with the carboxylate of Asp-60 (O−O distance of 2.8 Å), and the length of this hydrogen bond is the same as the one formed in the TRPS−**1** complex. The *o*-amino group of **3** also makes one hydrogen bond with the same carboxylate oxygen (N−O distance of 2.9 Å). Compounds **4** and **5** possess two unique atoms that were designed to enhance interactions with TRPS. The *p*-fluorine substituent of the ring in **4** does not participate in any polar interactions and is in proximity only to the CD1 carbon of Ile-153 (F−C distance of 3.1 Å). The sulfoxide oxygen of **5** seems to make a strong hydrogen bond with the hydroxyl group of Tyr-175 (O−O distance of 2.6 Å).

*Do Low-Barrier Hydrogen Bonds Exist in the Enzyme*−*Inhibitor Complexes?* In the complexes of TRPS with inhibitors **1** and **3**−**5**, the distance between one of the oxygens of the phosphonate group and the hydroxyl oxygen of Ser-235 has been refined to values of ≤2.5 Å, implying the involvement of a strong, low-barrier hydrogen bond (LBHB) in the stabilization of the enzyme−inhibitor complexes. The specific distance of this hydrogen bond for each of the inhibitors is as follows: **1**, 2.4 Å; **2**, 2.6 Å; **3**, 2.5 Å; **4**, 2.5 Å; and **5**, 2.5 Å. While the resolution of our data and the estimated coordinate errors for the derived models do not allow us to ascribe statistical significance to these very short distances, their presence in four out of five of our inhibitors supports further studies for verifying their significance. Confirmation of a low-barrier hydrogen bond will require NMR experiments to determine whether low-field proton chemical shifts and low fractionation factors are present.

The existence of LBHBs within enzyme active sites is controversial in light of recent theoretical (molecular mechanics and ab initio quantum mechanical) calculations (*36*, *37*) and NMR spectroscopic data (*38*). Nonetheless, small molecule and macromolecular crystallography at very high (atomic) resolution have confirmed the validity of very short hydrogen bonds. These bonds have been observed in crystal structures of salts of carboxylic and phosphorus-containing acids (*39*−*41*). In the case of phosphate crystals, $\pi$ bonding (of the $d_\pi-p_\pi$ type) within the phosphoryl group with its stronger potential acidity has been implicated as being responsible for a reduction of the effective radius of the oxygen atom, thus leading to a shortening of otherwise normal hydrogen bonds (*39*). Again, however, there is no general agreement (in the small molecule literature) for the cutoff value of the distance between the two heteroatoms that can distinguish these bonds from ordinary (high-barrier, double-welled) hydrogen bonds (*42*, *43*). There is a single case of a protein−ligand complex (phosphate complexed to phosphate binding protein) refined at 0.98 Å resolution in which a very short hydrogen bond is unambiguously observed (*44*). In this example, coordinate errors can be derived rigorously for all atoms by inversion of the final least-squares matrix, and are estimated to be 0.007 Å for the heteroatoms involved in the hydrogen bond. Interestingly, this bond was also observed in a lower-resolution structure (*45*). Very short hydrogen bonds have also been observed in a number of structures at resolutions varying between 1.6 and 2.3 Å involving phosphonate or phosphinate transition-state analogues complexed to carboxypeptidase A (*46*, *47*), thermolysin (*48*, *49*), penicillopepsin (*50*), HIV-1 protease (*51*), and endothiapepsin (*52*). In all of these complexes, one of the oxygens of the phosphorus-containing groups is shown to interact with one of the carboxylate oxygens of either a glutamic acid or an aspartic acid residue. The hydrogen bond distances (O···O distances) range between 2.2 and 2.5 Å. It has been proposed that such short, very strong, low-barrier hydrogen bonds can make a significant contribution to enzymatic catalysis and can have energies that are roughly 10 times higher than those of ordinary H-bonds (*32*−*35*, *53*).

In addition to very short distances, other criteria are also important for the establishment of low-barrier hydrogen bonds: (1) similar p$K_a$ values for the functional groups involved in the hydrogen bond and/or (2) the presence of permanent charge(s) in one of the two participating groups. The presence of a charge on the phosphonate group is supported by the lack of activity of analogues that are uncharged.[2] The participation of charges in low-barrier hydrogen bonds is particularly important for intermolecular interactions (*54*). The bond becomes stronger with the increasing strength of the electrical monopole. Formation of these bonds should be favored in sterically crowded active sites where some of the energy of the bond is used for the relief of steric strain, as has been recently proposed for serine proteases (*55*). However, we should note that it is not possible to isolate the contributions of the charge on the phosphonate to charge−charge interactions with the helix dipole or to the LBHB.

There are two other examples of very short hydrogen bonds in enzyme−ligand complexes that bear the closest chemical resemblance to the ones observed in our structures. In the complex of cytidine deaminase with a TSA inhibitor, an interaction occurs between an alcoholic hydroxyl of the inhibitor and a glutamate carboxylate oxygen with a refined O···O interatomic distance of 2.4 Å (*56*). A very similar bond between an aspartate carboxylate group and a hydroxyl of a sugar moiety of a trisaccharide is also found in the structure of a lysozyme−trisaccharide complex (*57*). In the

case of cytidine deaminase, the hydroxyl group is the predominant feature that distinguishes the transition state from the ground state of the substrate cytidine. In the case of lysozyme, however, this particular hydrogen bond is observed at a site far from where cleavage of the glycosidic bond of the sugar is proposed to occur. Thus, it may simply confer higher affinity of the ligand for the enzyme. In the case of tryptophan synthase, the hydroxyl of Ser-235 is playing a similar role in its interaction with the phosphonyl oxygen. To the best of our knowledge, this is the first time that such a strong hydrogen bond between a phosphonyl oxygen and an alcoholic hydroxyl oxygen is observed in enzyme−inhibitor complexes.

*Comparison of Inhibitor, Substrate (IGP), and Substrate Analogue (IPP and FIPP) Binding.* In all of our complexed structures, the inhibitors and protein active site residues superimpose fairly well (Figure 5A). The major differences occur with the alkyl carbon atoms that link the aromatic and phosphonate groups, particularly for inhibitors **3** and **4**. This difference is larger for **3**. The overall positioning of its alkyl chain is associated with an altered orientation of the phosphonate group relative to the other inhibitors. The altered position results in a weaker interaction between the phosphonate of **3** and the backbone amide of Gly-184 (O2−N distance of 3.5 Å compared to 2.9−3.0 Å for the other inhibitors).

Structures of the substrate IGP (*9*) and of the substrate analogues IPP (*6*) and FIPP (*11*) complexed to TRPS have been reported. Most of the interactions of the phosphonate group and the ortho substituent of the phenyl ring of our inhibitors are similar to those of the phosphate group and the indole nitrogen, respectively, of IGP in the TRPS−IGP complex (Figure 5B). However, the various protein−ligand complexes differ significantly in the conformations of active site residues Glu-49 and Phe-212. Glu-49 points away from the ligand in the complexes of TRPS with the phosphonate inhibitors (this study), IPP (*6*, *8*) or FIPP (*11*), whereas it is in an extended conformation in the free enzyme (*6*) and in the complex of the αD60N mutant with the natural substrate IGP (*9*). The extended conformation of Glu-49 allows it to form a hydrogen bond with IGP. Surprisingly, Glu-49 forms no apparent polar interactions in the active site of the free enzyme. Modeling of the phosphonate inhibitors, IPP, or FIPP in the active site of the free enzyme shows no steric clashes with Glu-49 (with the exception of the sulfoxide oxygen of **5** which would clash at distances of 2.2 and 2.1 Å from the side chain oxygens of Glu-49). The major reason for the different conformation of Glu-49 appears to be the loss of enthalpic stabilization provided by a hydrogen bond between the carboxylate of Glu-49 and the hydroxyl of C3′ of the glyceryl group of IGP. This in turn arises from the lack of a hydrogen bond donor at the C4 position of our inhibitors (equivalent to the C3′ of IGP) and also at the C3′ position of IPP and FIPP. The conformation of Glu-49 with the phosphonate inhibitors in the active site is stabilized by a new hydrogen bond network formed by Tyr-173, Tyr-104, and two water molecules. Differences are also observed in the conformation of Phe-212. The side chain of Phe-212 packs against the phenyl ring of the phosphonates and the indole rings of FIPP and IPP but adopts a different position in the free enzyme and the IGP complex, where it points away from the active site and becomes a surface residue.

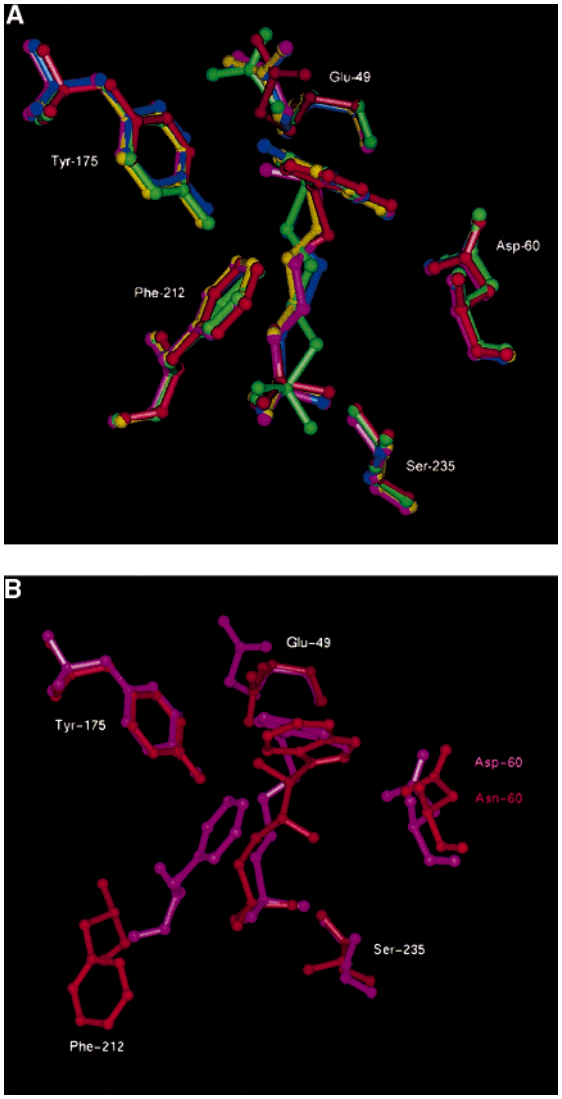


FIGURE 5: Structural comparisons of α-site ligands of tryptophan synthase. (A) Relative location of α-site ligands obtained by superposition of the α-subunits of each phosphonate−enzyme complex. The superposition shows that the relative orientation of active site residues is the same for each inhibitor: **1** (red), **2** (yellow), **3** (green), **4** (blue), and **5** (purple). (B) Superposition of the α-subunit active site residues of the αD60N tryptophan synthase mutant complexed with indole 3-glycerol phosphate and the wild-type TRPS complexed with a representative inhibitor (**1**) from our study: IGP (red) and **1** (purple).

Furthermore, if Phe-212 had the same orientation in our structures as in the free enzyme and IGP complex, it would sterically clash with Arg-179 (this residue is not visible in the free enzyme and IGP complex crystal structures). Therefore, the change in the conformation of Phe-212 stabilizes TRPS complex formation with the phosphonate inhibitors through hydrophobic and van der Waals interactions.

*Conformation of Flexible Loop αL6.* In contrast to the conclusion drawn by Rhee et al. (*8*) that both the α- and *β*-subunit active sites must be occupied with substrate to induce closure of loop αL6 (residues 177−191 in the α-subunit), and in agreement with the results obtained by Schneider et al. (*11*), we observe that the presence of an α-site ligand alone is sufficient for the ordering and closure of loop αL6 at the top of the α-site. Most of the residues of loop αL6 are visible in our complex structures (Figure 6).

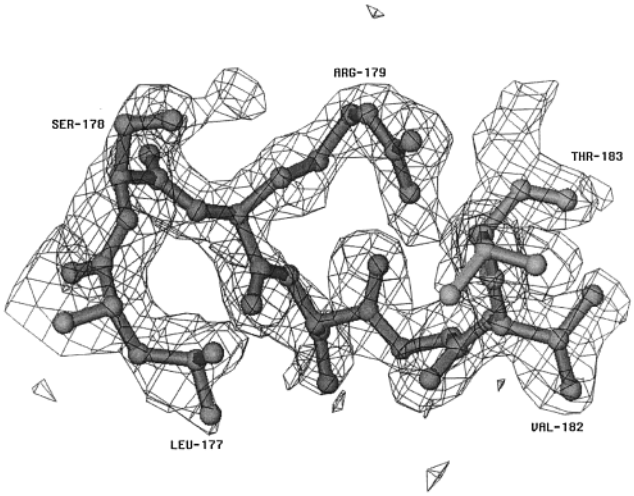


FIGURE 6: $2F_{obs} - F_{calc}$ $\sigma_A$-weighted simulated annealing omit electron density map for residues α177−α183 of loop αL6 of tryptophan synthase contoured at 1.3$\sigma$ showing the continuous density of the loop.

In particular, there is clear density for residues 177−183 which also have lower *B* factors compared to those of the remaining portion of the loop. Residues 190 and 191 remain unobserved in our complexes, although some electron density exists between residues α189 and α192. Gly-184 is the only loop residue that forms a hydrogen bond (involving its backbone nitrogen) with one of the phosphonate oxygens of the inhibitors (N−O distances of 2.9−3.0 Å). Furthermore, in this conformation, Thr-183 stabilizes Asp-60 by forming a hydrogen bond with one of its carboxylate oxygens.

*Implications for the Mechanism of Catalysis.* The transition state of the α-reaction is presumed to involve a tetrahedral carbon atom (Figure 1). The $sp^3$-hybridized sulfur atom in the inhibitors is meant to mimic this transition state. Therefore, analysis of the interactions between the inhibitors and the enzyme could be useful in understanding the catalytic mechanism. The C−S−C angle in all of the arylthioalkylphosphonate inhibitors in this study varies between 105° and 113°, which is close to the expected value for a tetrahedrally coordinated atom (109° 28′). This implies that the sulfur atom mimics the putative tetrahedral carbon atom in the transition state. It is interesting to note that among all of the inhibitor atoms, the position of the sulfur atom is most dissimilar from the corresponding $sp^2$-hybridized atoms of IGP, IPP, and FIPP (Figure 5B). The most likely explanation is the absence of ring constraints for our inhibitors, but substantial movement at this position as substrate is converted to product cannot be ruled out.

The transition state in the α-subunit active site is formed with the assistance of three functional groups: $B_1H$, $B_2$, and $B_3$. Asp-60 and Glu-49 have been previously identified as $B_2$ and $B_3$, respectively, but the identity of $B_1H$ has remained unknown (*9*). The structures presented here reinforce the idea that Asp-60 plays a catalytically important role as a base ($B_2$) that abstracts the proton from the indole nitrogen (NH) and facilitates indolenine tautomerization of IGP. In all of the complexes, the ortho substituent of the phenyl ring, which is in a position equivalent to that of the NH of indole and exerts similar electronic effects on the ring, interacts with the carboxylate of this particular aspartate residue. Our inhibitors do not possess any polar substituent (H-bond donor) on the C4 of the alkyl group, which is equivalent to the C3′ of the indole of IGP. Such a group could potentially mimic the interactions of the C3′ hydroxyl of IGP. Its absence from our inhibitors prevents us from confidently drawing conclusions based on these structures with respect to the nature of base $B_3$. However, the recently determined structure of the complex of a αD60N mutant of TRPS with the natural substrate IGP (*9*) revealed a strong hydrogen bond between one of the carboxylate oxygens of Glu-49 and the C-3′ hydroxyl of IGP, implying that this group can serve as a base that will deprotonate the C3′ hydroxyl during catalysis and facilitate IGP cleavage.

The structures in this study cast doubt on the proposed role of Glu-49 as the acid ($B_1H$) that protonates the C3 of indole. The carboxylate of Glu-49 points away from the sulfur atom (which mimics C3 of indole) in all of our inhibitors. The distance between the sulfur atoms and the carboxylate oxygen of Glu-49 is refined to values ranging between 4.4 and 6.0 Å. This could imply that (i) another functional group participates in protonation of C3, (ii) our inhibitors do not represent an appropriate model of the transition state of the α-reaction, or (iii) close proximity between the carboxylate of Glu-49 and C3 (or S) is achieved only transiently in the active site and is unobservable in our time-averaged diffraction experiment. The enzyme−inhibitor structures here highlight the possible role that Tyr-175 can play in this part of the α-reaction. In particular, the distance between the hydroxyl of Tyr-175 and the sulfur, which mimics the C3 atom that becomes protonated during the reaction, varies between 3.5 and 3.9 Å in all of the inhibitor structures, and is significantly shorter than the corresponding distance from the carboxylate of Glu-49. It is interesting that Tyr-175 is one of the sites of missense mutations that inactivates the α-subunit (*58*, *59*).

The significant catalytic activity reported for the αY175F mutant suggests that Tyr-175 is not the acid catalyst (*18*). Nagata et al. conclude that the only other functional group that could fulfill this role is Glu-49. However, other conclusions regarding the catalytic roles of these residues are also possible. For example, the phenolic hydroxyl of Tyr-175 may function as the catalytic acid in the wild-type enzyme. Upon replacement with phenylalanine, the mutant enzyme may undergo a change in mechanism that now allows Glu-49 or a water molecule that is sufficiently polarized in the low-dielectric environment of the active site to compensate for the absence of Tyr-175 and function as the catalytic acid. Such a change in catalytic mechanism has been observed in a mutant of triosephosphate isomerase with the replacement of a catalytically important electrophilic group in the active site (*60*). The absence of detailed structural and mechanistic information for the Y175F mutant does not allow the significant catalytic activity observed for the α-reaction to be interpreted with confidence. Additional studies of the Y175F mutant are required to resolve this issue.

*Design of More Potent Inhibitors*. The design of indolenine-based compounds may also produce potent inhibitors of TRPS. This idea stems directly from the proposed mechanism for the α-reaction of TRPS (Figure 1). Formation of an indolenine tautomer is suggested to facilitate the C−C bond cleavage in IGP. It is noteworthy that in an effort to study the mechanism of the α-reaction, indoline derivatives (2,3-dihydroindoles that contain a tetrahedral carbon at C3)

were synthesized and shown to be strong competitive inhibitors of the tryptophan synthase reaction (*61*, *62*). These compounds, however, exhibited 10−100-fold weaker activity than the phosphonate inhibitors used in the study presented here. This could be due to the fact that the indoline compounds, oxindolyl-L-alanine and 2,3-dihydro-L-tryptophan, contain an aminocarboxyethyl instead of the phosphonylbut(en)yl side chain of our inhibitors. Presumably, the indoline analogues cannot form electrostatic and hydrophobic interactions with active site residues that are as potent as those of our inhibitors.

Examination of the binding site for arylthiolphosphonate inhibitors in this study or previously published inhibitors of tryptophan synthase reveals additional interactions that could be exploited to improve affinity. For example, in many of the inhibitors, favorable van der Waals contacts with Ile-64, Thr-183, and Gly-234 could be made by replacing hydrogen atoms along the alkyl chains with larger atoms or functional groups. The actual positions that should be substituted are different for each inhibitor due to subtle differences in binding orientations between each inhibitor. The differences in binding among the ligands of tryptophan synthase that impact analogue design are best illustrated by the arylthiolphosphonate inhibitors of this study. The C4 atom of the arylthiolphosphonates should mimic the C3′ atom of the natural substrate to which a hydroxyl group is attached. However, placement of a hydroxyl group at C4 of these inhibitors is unlikely to exploit the potential interaction with Glu-49 due to stereochemical differences that result from the different position of sulfur relative to the C3 position of the indole ring of the substrate and substrate analogues. Indoline analogues may be more appropriate for hydroxyl modification at this position because the constraints imposed by the five-membered ring would allow the hydroxyl to occupy a position similar to that of IGP. Incorporating features from indoline rings with the phosphonate-based side chains used in the study presented here may result in further improvements in the potency of inhibitors for TRPS and lead to effective antibacterial, antifungal, and herbicidal agents.

## ACKNOWLEDGMENT

We thank Paul Pepin and Amy Jan for help with X-ray data collection.

## REFERENCES

1. Bornemann, S., Lowe, D. J., and Thorneley, R. N. (1996) *Biochem. Soc. Trans. 24*, 84−88.
2. Kishore, G. M., and Shah, D. M. (1988) *Annu. Rev. Biochem. 57*, 627−663.
3. Miles, E. W. (1991) *Adv. Enzymol. Relat. Areas Mol. Biol. 64*, 93−172.
4. Miles, E. W. (1995) *Subcell. Biochem. 24*, 207−254.
5. Pan, P., Woehl, E., and Dunn, M. F. (1997) *Trends Biochem. Sci. 22*, 22−27.
6. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., and Davies, D. R. (1988) *J. Biol. Chem. 263*, 17857−17871.
7. Rhee, S., Parris, K. D., Ahmed, S. A., Miles, E. W., and Davies, D. R. (1996) *Biochemistry 35*, 4211−4221.
8. Rhee, S., Parris, K. D., Hyde, C. C., Ahmed, S. A., Miles, E. W., and Davies, D. R. (1997) *Biochemistry 36*, 7664−7680.
9. Rhee, S., Miles, E. W., and Davies, D. R. (1998) *J. Biol. Chem. 273*, 8553−8555.
10. Schlichting, I., Yang, X. J., Miles, E. W., Kim, A. Y., and Anderson, K. S. (1994) *J. Biol. Chem. 269*, 26591−26593.
11. Schneider, T. R., Gerhardt, E., Lee, M., Liang, P. H., Anderson, K. S., and Schlichting, I. (1998) *Biochemistry 37*, 5394−5406.
12. Anderson, K. S., Kim, A. Y., Quillen, J. M., Sayers, E., Yang, X. J., and Miles, E. W. (1995) *J. Biol. Chem. 270*, 29936−29944.
13. Anderson, K. S., Miles, E. W., and Johnson, K. A. (1991) *J. Biol. Chem. 266*, 8020−8033.
14. Brzovic, P. S., Ngo, K. N., and Dunn, M. F. (1992) *Biochemistry 31*, 3831−3839.
15. Leja, C. A., Woehl, E. U., and Dunn, M. F. (1995) *Biochemistry 34*, 6552−6561.
16. Walsh, C. T. (1979) *Enzymatic Reaction Mechanisms*, Freeman and Co., San Francisco, CA.
17. Miles, E. W., McPhie, P., and Yutani, K. (1988) *J. Biol. Chem. 263*, 8611−8614.
18. Nagata, S., Hyde, C. C., and Miles, E. W. (1989) *J. Biol. Chem. 264*, 6288−6296.
19. Miles, E. W., Kawasaki, H., Ahmed, S. A., Morita, H., Morita, H., and Nagata, S. (1989) *J. Biol. Chem. 264*, 6280−6287.
20. Creighton, T. E., and Yanofsky, C. (1966) *J. Biol. Chem. 241*, 980−990.
21. Kirschner, K., Wiskocil, R. L., Foehn, M., and Rezeau, L. (1975) *Eur. J. Biochem. 60*, 513−523.
22. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−325.
23. Dodson, E. J., Winn, M., and Ralph, A. (1997) *Methods Enzymol. 277*, 620−633.
24. Brunger, A. T. (1997) *X-PLOR*, version 3.851, Yale University Press, New Haven, CT.
25. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
26. Brunger, A. T. (1992) *Nature 355*, 472−475.
27. Jiang, J. S., and Brunger, A. T. (1994) *J. Mol. Biol. 243*, 100−115.
28. Luzzati, V. (1952) *Acta Crystallogr. 5*, 802−810.
29. Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
30. Read, R. J. (1990) *Acta Crystallogr. A46*, 900−912.
31. Hodel, A., Kim, S.-H., and Brunger, A. T. (1992) *Acta Crystallogr. A48*, 851−858.
32. Cleland, W. W. (1992) *Biochemistry 31*, 317−319.
33. Cleland, W. W., and Kreevoy, M. M. (1994) *Science 264*, 1887−1890.
34. Gerlt, J. A., and Gassman, P. G. (1993) *J. Am. Chem. Soc. 115*, 11552−11568.
35. Gerlt, J. A., and Gassman, P. G. (1993) *Biochemistry 32*, 11943−11952.
36. Scheiner, S., and Kar, T. (1995) *J. Am. Chem. Soc. 117*, 6970−6975.
37. Warshel, A., and Papazyan, A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 13665−13670.
38. Ash, E. L., Sudmeier, J. L., De Fabo, E. C., and Bachovchin, W. W. (1997) *Science 278*, 1128−1132.
39. Calleri, M., and Speakman, J. C. (1964) *Acta Crystallogr. 17*, 1097−1103.
40. Darlow, S. F., and Cochran, W. (1961) *Acta Crystallogr. 14*, 1256−1260.
41. Speakman, J. C. (1972) *Struct. Bonding 12*, 141−192.
42. Hibbert, F., and Emsley, J. (1990) *Adv. Phys. Org. Chem. 48*, 255−347.
43. Perrin, C. L., and Nielson, J. B. (1997) *Annu. Rev. Phys. Chem. 48*, 511−544.
44. Wang, Z., Luecke, H., Yao, N., and Quiocho, F. A. (1997) *Nat. Struct. Biol. 4*, 519−522.
45. Luecke, H., and Quiocho, F. A. (1990) *Nature 347*, 402−406.
46. Kim, H., and Lipscomb, W. N. (1990) *Biochemistry 29*, 5546−5555.
47. Kim, H., and Lipscomb, W. N. (1991) *Biochemistry 30*, 8171−8180.
48. Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H., and Matthews, B. W. (1987) *Biochemistry 26*, 8542−8553.
49. Tronrud, D. E., Monzingo, A. F., and Matthews, B. W. (1986) *Eur. J. Biochem. 157*, 261−268.

50. Fraser, M. E., Strynadka, N. C. J., Bartlett, P. A., Hanson, J. E. J., and James, M. N. G. (1992) *Biochemistry 31*, 5201−5214.
51. Abdel-Meguid, S. S., Zhao, B., Murthy, K. H. M., Winborne, E., Choi, J.-K., DesJarlais, R. L., Minnich, M. D., Culp, J. S., Debouck, C., Tomaszek, T. A., Jr., Meek, T. D., and Dreyer, G. B. (1993) *Biochemistry 32*, 7972−7980.
52. Dealwis, C. (1993) Ph.D. Thesis, Birkbeck College, London, U.K.
53. Frey, P. A., Whitt, S. A., and Tobin, J. B. (1994) *Science 264*, 1927−1930.
54. Jeffrey, G. A., and Saenger, W. (1991) *Hydrogen Bonding in Biological Structures*, Springer-Verlag, Berlin.
55. Cassidy, C. S., Lin, J., and Frey, P. A. (1997) *Biochemistry 36*, 4576−4584.
56. Xiang, S., Short, S. A., Wolfenden, R., and Carter, C. W., Jr. (1995) *Biochemistry 34*, 4516−4523.
57. Strynadka, N. C. J., and James, M. N. G. (1991) *J. Mol. Biol. 220*, 401−424.
58. Yanofsky, C., and Horn, V. (1972) *J. Biol. Chem. 247*, 4494−4498.
59. Helinski, D. E., and Yanofsky, C. (1963) *J. Biol. Chem. 238*, 1043−1048.
60. Nickbarg, E. B., Davenport, R. C., Petsko, G. A., and Knowles, J. R. (1988) *Biochemistry 27*, 5948−5960.
61. Phillips, R. S., Miles, E. W., and Cohen, L. A. (1984) *Biochemistry 23*, 6228−6234.
62. Phillips, R. S., Miles, E. W., and Cohen, L. A. (1985) *J. Biol. Chem. 260*, 14665−14670.
63. Finn, J., Langevine, C., Birk, I., Birk, J., Nickerson, K., and Rodaway, S. (1999) *Bioorg. Med. Chem. Lett. 9*, 2297−2302.

BI9907734