Crystal Structure and Mechanism of Tryptophan 2,3-Dioxygenase, a Heme Enzyme Involved in Tryptophan Catabolism and in Quinolinate Biosynthesis[†,‡]

Yang Zhang, Seong A. Kang, Tathagata Mukherjee, Shridhar Bale, Brian R. Crane, Tadhg P. Begley,* and Steven E. Ealick*

Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853

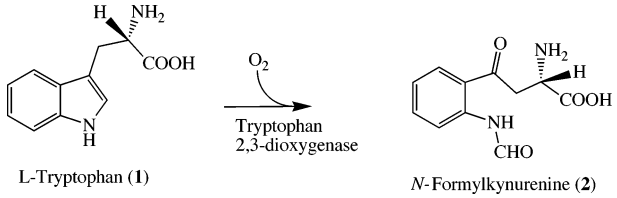
Received September 27, 2006

ABSTRACT: The structure of tryptophan 2,3-dioxygenase (TDO) from *Ralstonia metallidurans* was determined at 2.4 Å. TDO catalyzes the irreversible oxidation of L-tryptophan to *N*-formyl kynurenine, which is the initial step in tryptophan catabolism. TDO is a heme-containing enzyme and is highly specific for its substrate L-tryptophan. The structure is a tetramer with a heme cofactor bound at each active site. The monomeric fold, as well as the heme binding site, is similar to that of the large domain of indoleamine 2,3-dioxygenase, an enzyme that catalyzes the same reaction except with a broader substrate tolerance. Modeling of the putative (*S*)-tryptophan hydroperoxide intermediate into the active site, as well as substrate analogue and mutagenesis studies, are consistent with a Criegee mechanism for the reaction.

Tryptophan 2,3-dioxygenase (TDO)[1] catalyzes the initial step in tryptophan catabolism. This pathway also generates quinolinate, which is the precursor to the pyridine ring of the essential redox cofactor NAD. The reaction converts L-tryptophan (**1**, L-Trp), the least abundant essential amino acid found in mammals, to *N*-formyl kynurenine (**2**, Scheme 1) by cleaving the 2,3 bond of L-Trp and incorporating both atoms of molecular oxygen (*1*, *2*). The majority of dietary tryptophan not used in protein synthesis is metabolized by this pathway, while a small amount of tryptophan is utilized to synthesize the neurotransmitter serotonin, which can be further converted to melatonin in the brain pineal body (*3*).

TDO was first identified in the liver of mammals by Kotake and Masayama in 1936 (*2*). Although enzyme activity has also been found in the skin (*4*) and mouse early concepti (*5*), TDO is primarily expressed in liver but not other tissues. Interestingly, another enzyme capable of catalyzing the same oxidative ring opening of D-tryptophan (D-Trp) was discovered in 1967 by Higuchi, Yamamoto, and Hayaishi (*6, 7*). While TDO is specific for L-tryptophan, this later discovered enzyme exhibited a much broader substrate acceptance, including L-Trp, D-Trp, 5-hydroxy-Trp, tryptamine, and serotonin (*8*−*10*), and was therefore named indoleamine 2,3-dioxygenase (IDO). IDO also has a much wider distribution. The enzyme has been found in most tissues except liver, including stomach, intestines, colon, kidney, spleen, lung, and brain (*3*, *11*). Despite the similarity of the reactions catalyzed, the two enzymes share little sequence identity.

Scheme 1



L-Tryptophan (**1**) *N*-Formylkynurenine (**2**)

While the chemical mechanism is still poorly understood for both enzymes, a recent breakthrough is the structure determination of human IDO. This protein binds heme at the interface of two domains (*12*, *13*). In addition, the Northeast Structural Genomics Consortium recently determined the structure of heme-free IDO from *Shewanella oneidensis* (PDB code: 1ZEE). Human IDO and *S. oneidensis* IDO are structurally homologous even though the sequence identity is only ∼14%.

The tryptophan-based quinolinate biosynthesis, also called the kynurenine pathway, had been generally believed unique to the eukaryotes, and a different pathway using aspartate and dihydroxyacetone to produce quinolinate was found in prokaryotes. However, the enzyme activities involved in the kynurenine pathway were also identified in several bacteria (*14*). The bacterial TDOs share approximately 20−30% sequence identity with their vertebrate orthologs. The structure of TDO was first determined from *Xanthomonas campestris* (PDB code: 1YW0); however, the heme group was not present, and the active site was not identified.

Products of tryptophan oxidation have diverse physiological effects on the immune and nervous systems. In particular, IDO suppresses T-cell-mediated immune responses, and quinolinic acid, which is derived from tryptophan, is a neurotoxin (*15*−*18*). Tryptophan oxidation is therefore of considerable biomedical significance. Increased understanding of the mechanism of this oxidation will facilitate the design of inhibitors.

[†] This work was supported by National Institutes of Health Grants GM069618 (to T.P.B. and S.E.E.). S.E.E. is indebted to the W. M. Keck Foundation and the Lucille P. Markey Charitable Trust.

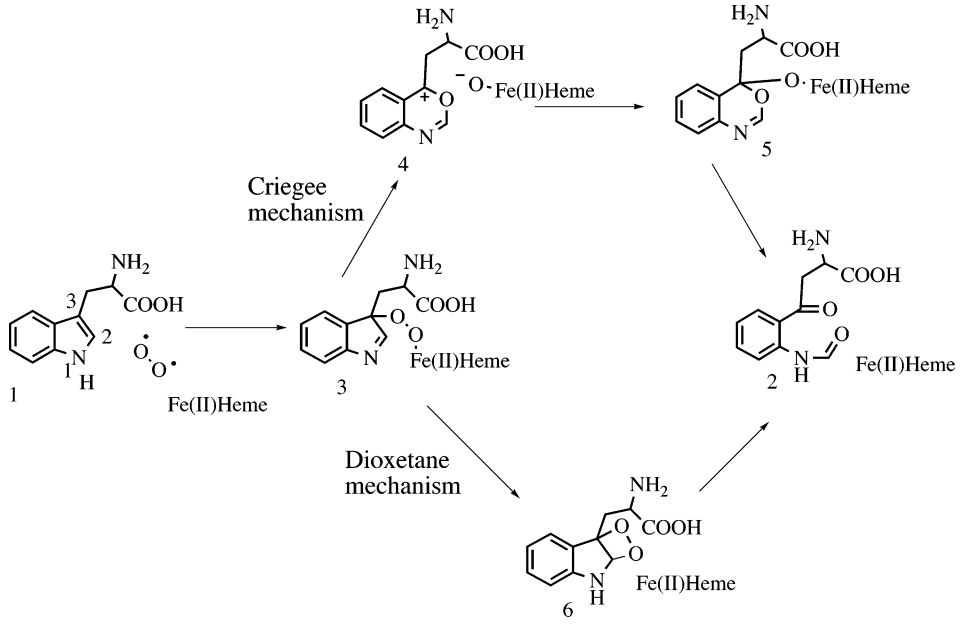
[‡] The Brookhaven Protein Data Bank code is 2NOX.

* To whom correspondence should be addressed. Telephone: (607) 255-7961. Fax: (607) 255-1227. E-mail: see3@cornell.edu or tpb2@cornell.edu.

[1] Abbreviations: tryptophan 2,3-dioxygenase, TDO; L-tryptophan, L-Trp; indoleamine 2,3-dioxygenase, IDO; D-tryptophan, D-Trp; noncrystallographic symmetry, NCS.

10.1021/bi0620095 CCC: $37.00
Published on Web 12/08/2006

Scheme 2



In this paper, we report the structure of TDO from *Ralstonia metallidurans* with cofactor heme bound in the active site. The structure was determined by molecular replacement using the *X. campestris* structure as the search model. The *R. metallidurans* TDO crystal structure contains four tetramers per asymmetric unit. Each homotetramer consists of 299 amino acids per monomer, corresponding to a total molecular mass of 134 kDa. The tetramer exhibits 222 symmetry, and the monomeric fold of TDO resembles the large domain of IDO. Each active site resides at the interface of two monomers, comprised mostly by one monomer and the N-terminus of the other monomer. A type-b heme molecule is bound at each active site. Modeling and mutagenesis studies were carried out to investigate the enzyme mechanism (Scheme 2).

MATERIALS AND METHODS

Supplies. Antibiotics, 5-aminolevulinic acid hydrochloride, and the buffer components were purchased from Sigma-Aldrich. Plasmid purification kits and Ni-NTA resin were obtained from Qiagen (Valencia, CA). Methanol D4 was procured from Cambridge Isotope Laboratories, Inc. (CIL).

Construction, Expression, and Purification of Ralstonia metallidurans TDO. The genes encoding *R. metallidurans* TDO were PCR cloned into vector pET28a (Novagen) and expressed with a fused 6-His tag in *Escherichia coli* strain BL21 (DE3) (Novagen) in Terrific Broth (DIFCO) with kanamycin selection (25 μg/mL). A heme precursor, δ-aminolevulinic acid, was required for the expression of the fully active enzyme and was added to the cell culture at 20 μg/mL prior to induction. To encourage crystal formation, TDO was expressed from a series of constructs that truncated the N-terminus of the original TDO gene. We found a dramatic effect of N-terminal start positions on protein solubility and stability. Diffraction quality TDO crystals were grown from a construct that lacked 18 N-terminal residues (Δ18TDO). The recombinant protein was purified by metal-chelate affinity chromatography using a Ni-NTA column (Novagen). Freshly purified protein (~25 mg/mL) in 200 mM imidazole, 300 mM sodium chloride, 50 mM sodium phosphate at pH 8.0 was used for the crystallization experiments immediately without further purification steps.

Table 1: Data Collection Statistics[a]

wavelength (Å)	1.1000
resolution (Å)	2.4
reflections	479205
unique reflections	183150
completeness (%)	97.5 (97.0)
R_{sym}[b] (%)	7.4 (49.8)
I/σ	13.7 (2.3)
redundancy	2.1 (2.1)

[a] Values for the highest resolution shell are given in parentheses. [b] $R_{sym} = \sum\sum_i|I_i - \langle I\rangle|/\sum\langle I\rangle$, where $\langle I\rangle$ is the mean intensity of the N reflections with intensities I_i and common indices *h,k,l*.

Crystallization of R. metallidurans TDO and X-ray Intensity Measurements. Initial conditions for growing TDO crystals were found using commercial screening solutions (Hampton Research) and improved for better diffraction quality after extensive optimization of conditions. Optimized crystals of Δ18TDO (25 mg/mL) grew by vapor diffusion against a reservoir of 0.8−1.2 M tri-sodium citrate dihydrate pH 6.5 in the presence of 1 mM L-Trp at 22 °C. Crystals of Δ18TDO belong to the space group *P*1 with unit cell dimensions of $a = 72.5$ Å, $b = 132.1$ Å, $c = 139.9$ Å, $\alpha = 67.0°$, $\beta = 85.1°$, and $\gamma = 89.9°$. X-ray diffraction data were collected at the National Synchrotron Light Source (NSLS), beamline X25. Temperature was controlled by an Oxford cryodevice, and 25% (vol/vol) glycerol cryoprotectant was used when necessary. All diffraction data were indexed, integrated, and scaled using HKL2000 (*19*). Data processing statistics are summarized in Table 1.

Structure Determination and Refinement. The structure of *R. metallidurans* TDO with the N-terminal 18 residues truncated was determined by molecular replacement using the CNS software package (*20*). There are four complete tetramers in the *P*1 unit cell, corresponding to a solvent content of 50.3% and Matthews coefficient of 2.47 Å^3/Da. A tetramer of *X. campestris* TDO (PDB code: 1YW0), which

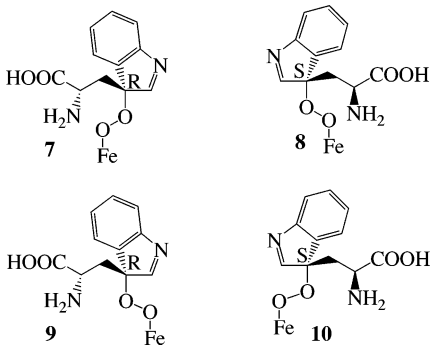
Table 2: Refinement Statistics

resolution (Å)	2.4
number of protein atoms	34130
number of water molecules	1543
root-mean-square deviation from ideal geometry	
bonds (Å)	0.012
angles (deg)	1.487
R factor[a] (%)	21.2
R_{free}[b] (%)	27.3
Ramachandran plot	
most favored regions (%)	92.3
additional allowed regions (%)	7.6
generously allowed regions (%)	0.1
disallowed regions (%)	0.0
average B factors (Å^2)	
protein residues	45.6
water molecules	44.3
heme	41.9

[a] R factor $= \sum_{hkl}||F_{obs}| - k|F_{cal}||/\sum_{hkl}|F_{obs}|$ where F_{obs} and F_{cal} are observed and calculated structure factors, respectively. [b] For R_{free} the sum is extended over a subset of reflections (5%) excluded from all stages of refinement.

shares about 47% sequence identity with *R. metallidurans* TDO but lacks the heme group, was used as a search probe to separately locate each of the four tetramers. The initial rigid body refinement and simulated annealing resulted in an *R*-factor of 37%. A composite omit map was calculated and averaged based on the 16-fold noncrystallographic symmetry (NCS) using the program RAVE (*21*). The averaged map showed clear density and allowed unambiguous positioning of the heme groups. The 16 iron positions were further confirmed by calculating an anomalous difference Fourier map using data collected at the iron edge (data not shown). Residues 139−142 and residues 277−282, which correspond to the missing residues 122−125 and 260−265 in the structure of *X. campestris* TDO, were built in based on the averaged density. Residues 268−276, which are also missing from the *X. campestris* TDO structure (residues 251−259), showed clear density in monomers A, D, H, I, L, N, and P, but various extents of flexibility in other monomers. Simulated annealing, *B*-factor refinement, and water placement were implemented with CNS (*20*) and Refmac5 (*22*), amidst rounds of rebuilding with the program O (*23*). The NCS restraints were applied throughout the refinement and gradually loosened up on the regions that exhibited variations among monomers. All residues of the final model have good stereochemistry with no disfavored backbone and dihedral angles. The refinement statistics are summarized in Table 2.

Modeling of the Tryptophan Hydroperoxide into the Active Site. The modeling of the putative hydroperoxide intermediate (**7**, **8**, **9**, and **10**) into the active site of TDO was carried out using the program Macromodel, version 7.2 (*24*). Modeling simulations were preformed in which the Cβ−Cγ torsion was either constrained (anti periplanar or syn coplanar) or allowed to freely rotate. The base shell of atoms included all residues within 20.0 Å of the heme iron and was used as the starting model for energy minimization. The intermediate was built into the active site, followed by removal of water molecules and hydrogen atoms were added appropriately. The starting structure was subjected to mixed Monte Carlo MCMM/lowmode conformational search steps allowing residues in the 4 Å shell around the active site to move. This included the key residues Tyr43, Phe68, His72, and Arg134. The conformational search allowed torsional rotation about the Fe−O1, O1−O2, O2−Cγ, Cγ−Cβ, Cβ−Cα, and Cα−C bonds (O1 and O2 are the two incorporated oxygen atoms). The generated structures were energy minimized to a gradient of 0.01 kJ/mol Å *in vacuo* using the AMBER* force field (*25*, *26*), an electrostatic potential with a distance dependent 4r dielectric treatment and a TNCG minimization technique (*27*). The AMBER* parameters for the heme were adapted from D. A. Giammona (*28*). To test the effect of solvation on the minimization, the conformational search was repeated with a minimization that uses the GB/SA (*29*) solvation model. This search yielded a global minimum similar to the one from the *in vacuo* search.



Mutagenesis of RmTDO. Standard methods were used for DNA manipulations (*30*, *31*). Plasmid DNA was purified with a Qiagen Miniprep kit. *E. coli* strain MachI (Invitrogen) was used as a recipient for transformations during plasmid construction and for plasmid propagation and storage.

Site-directed mutagenesis was performed on pRmTDO.PXHta by a standard PCR protocol using *Pfu*Turbo DNA polymerase per the manufacturer's instructions (Invitrogen) and *Dpn*I (New England Biolabs) to digest the methylated parental DNA prior to transformation. For each mutant, a third primer was designed to screen for the presence of the mutant by colony PCR with an appropriate vector-specific primer. Only clones that produced a PCR product were sequenced. In every case, the mutagenesis primer pair consisted of the primer whose sequence is in Table 3 and its reverse complement.

Activity of TDO Mutants. Substrate solutions of varying concentrations were prepared by dissolving the required amount of L-tryptophan and ascorbic acid (reducing agent to keep the enzyme in the Fe(II) oxidation state) in a molar ratio of 4:1 in 100 mM phosphate buffer at pH 7.0. 50 μL of freshly purified enzyme (strength varying from 100 to 200 μM depending on the preparation of enzyme) was added to 450 μL of the substrate solution and the apparent catalytic activity of all the mutants and wild type was determined by measuring the rate of product formation at 321 nm ($\epsilon = 3152$ M^{-1} cm^{-1} for NFK). The K_m for tryptophan and k_{cat} for the active mutants and wild type were determined by fitting the rate of product formation versus substrate concentration to the Michaelis−Menten equation using Grafit 5.0.11 (*32*).

RESULTS

Crystal Structure of R. metallidurans TDO. We have determined the structure of *R. metallidurans* TDO by molecular replacement at 2.4 Å in space group *P*1. Four tetramers (ABCD, EFGH, IJKL, and MNOP) were found in

Table 3: Primer Sequences and Used for Mutagenesis and Screening

mutant	mutagenesis primer (top strand)/screening primer
Y130F	GATGACGCCGCCCGAGTTCAGCGCGATGCGCCCATACCTGACGCCGCCCGAGTTCAGC
F68A	GATCACAATGAGATGCTCGCAATCGTCCAACACCAGACCACCGATCACAATGAGATGCTCGC
T271A	GGTGCCGCCCGCGCCGCGCTTGAATCCGATTACCCGCTCGACTGTGGCTGACGCCTTCGGTGCCGCCCGC
R134A	GAGTACTCGGCCATGGCTCCATACCTGGGGGCATCGTCCGCCCCCAGGTATGGAGC

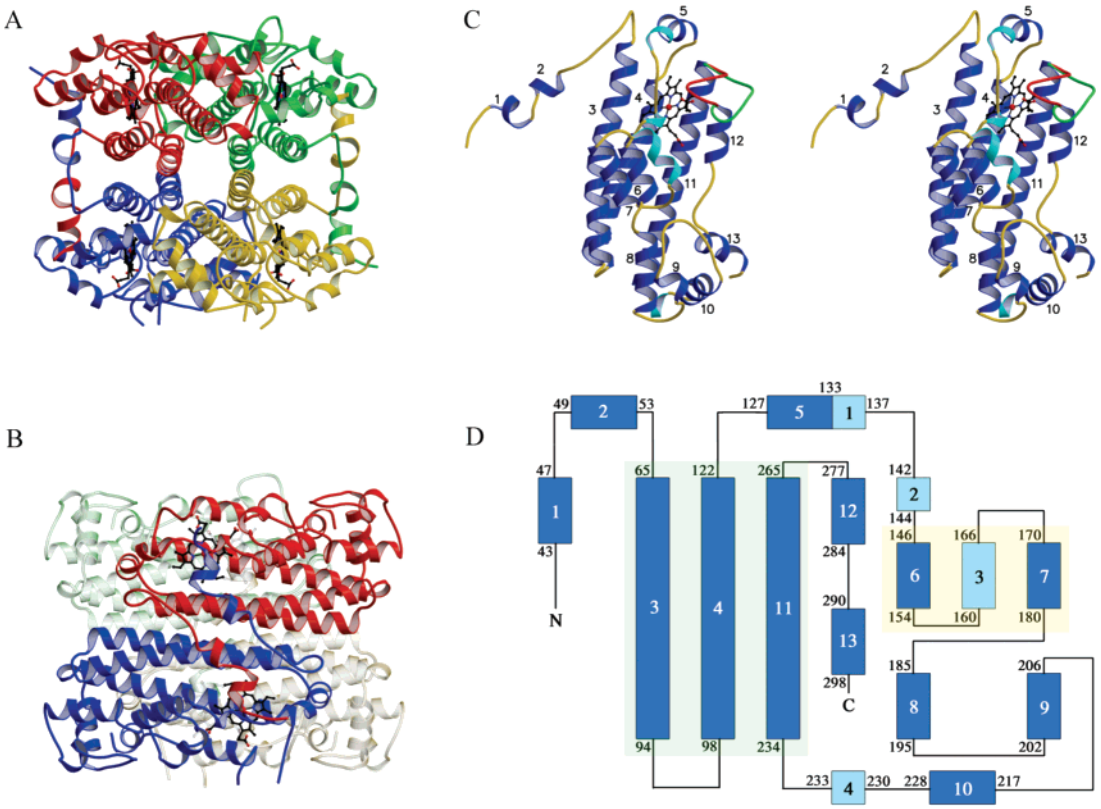


FIGURE 1: Structure of TDO. The TDO tetramer is color coded by subunits and shown down a 2-fold axis in panels (A) and (B). (C) Stereodiagram of TDO monomer shown with α-helices in dark blue and 3_{10} helices in light blue. Loop 266−276 adopts two conformations in different subunits. The open conformation is in green and the closed conformation is in red. (D) The topology diagram with the helical bundles highlighted.

the unit cell and were related by pseudo-four-fold noncrystallographic symmetry. The pseudo-four-fold axis is nearly parallel to the crystallographic *a*-axis. Each homotetramer has approximate dimensions of 70 × 70 × 60 Å. No systematic global differences were observed among the four tetramers; however, the 16 monomers showed significant variation at the N-terminus and the loop region comprised of residues 266−276. The overall three-dimensional structure of *R. metallidurans* TDO is presented in Figure 1.

In each tetramer, the four monomers are related by three mutually perpendicular two-fold axes (222 point symmetry). Each molecule has a central channel formed by long helices from each of the four monomers. The channel has a diameter of approximately 5 Å and is mostly hydrophilic. Using tetramer ABCD as a representative, adjacent monomers have more extensive interactions at the A−B interface than at the A−D interface, burying a surface area of 6200 and 2700 Å^2 at the two interfaces, respectively. The dimer (AB, CD) contains a heme group within one monomer, but the cavity for the predicted L-Trp binding site is partially formed by the adjacent monomer. Therefore, the tetramer can be viewed as a dimer of dimers, in which each dimer contains all of the amino acid residues required to form two complete active sites. The subunit interfaces are formed by a mixture of hydrophobic interactions, hydrogen-bond interactions, and salt bridges.

The TDO monomer is an all α-helical structure, consisting of 13 α-helices and four 3_{10} helices. The core motif of TDO is composed of three long α-helices (α3, α4, and α11), which also form most of the heme binding site. The three longest α-helices are flanked by the shorter α-helices and 3_{10}-helices. The interhelical interfaces are lined mostly with hydrophobic side chains. The unusually long (seven residues, two turns) helix 3_{10}-3 and helices α6 and α7 form a smaller three helix bundle near the heme binding site. The long helical bundle is surrounded by five other short α-helices (α5, α8, α9, α10, and α13). In addition, the two N-terminal α-helices from the adjacent monomer of the major dimer cross over the surface of the helical bundle and pack against α3 and loop 138−142.

Each monomer contains a heme binding site, which is located at the end of the long helical bundle and formed by six α-helices (α3, α4, α5, α6, α11, and α12), two 3_{10} helices (3_{10}-1 and 3_{10}-2), two glycine-rich loops (loop 138−142 and loop 266−276), and α1 from the adjacent subunit (Figure 2A). The heme binding site is described in detail below.

Glycine-Rich Loops and Alternate Conformations. The structure of *R. metallidurans* TDO has two glycine-rich loops

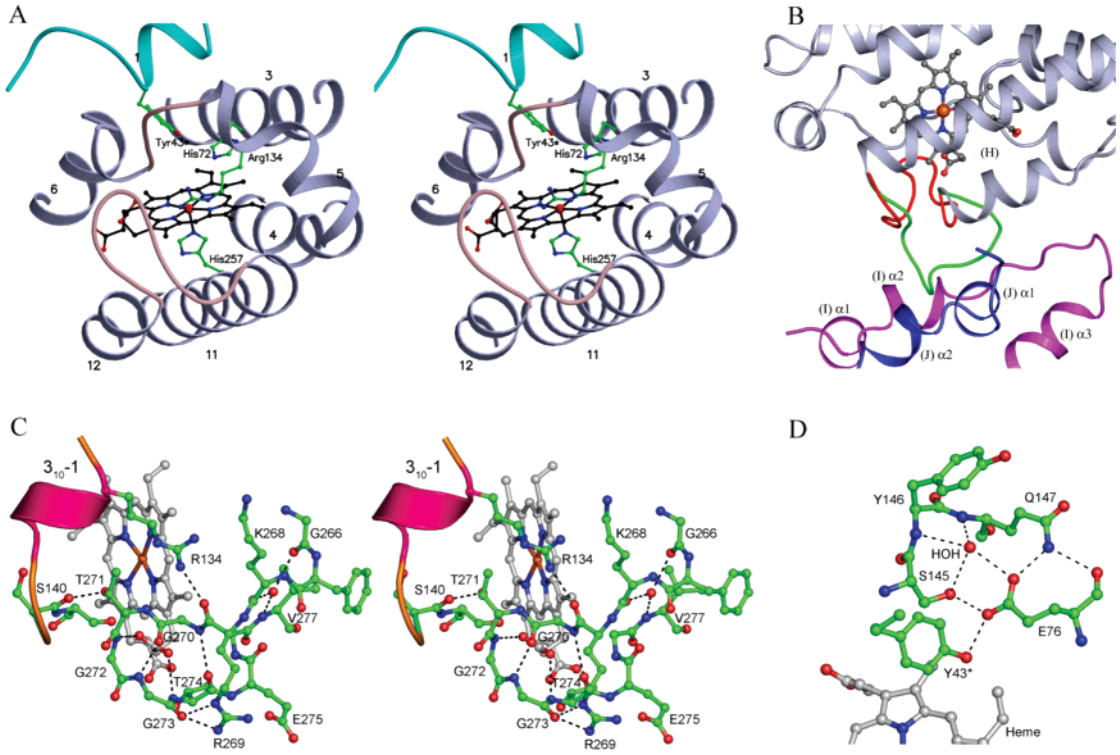


FIGURE 2: The open and closed conformations and the heme binding environment. (A) A heme group is bound in the closed conformation. Two glycine-rich loops are colored in pink. The N-terminal loop and α1 from the adjacent subunit are colored in cyan. (B) Loop 266−276 from subunit H (molecule EFGH) is in the open conformation (in green) and interacts with an adjacent molecule (IJKL). The closed conformation of monomer A is superimposed and colored in red. (C) Loop 266−275 from monomer A (closed form) is shown in ball-and-stick representation, and helix 3_{10}-1 and the loop 138−142 in ribbon. The hydrogen bonds are indicated by dashed lines. (D) A hydrogen-bond network is found near the heme group in both conformations. Three of the residues involved, Tyr43*, Glu76, and Ser145, are strictly conserved.

formed by residues 138−142 and 266−276, for which the corresponding residues in the *X. campestris* TDO structure are disordered. Both loops are involved in formation of the heme binding site. Loop 138−142 was built with a similar conformation in each of the 16 monomers; however, the loops showed relatively high-temperature factors after refinement. Loop 266−276, showed different extents of flexibility among monomers as reflected by the quality of the electron density and the magnitudes of the refined temperature factors. The loop was included in monomers A, D, H, I, L, N, and P but omitted or built partially for the other monomers. The seven monomers for which the entire loop 266−276 could be built show the loop to reside primarily in two different conformations (Table 4). Monomers A and I show a closed conformation relative to the heme group, while monomers D, H, L, N, and P show an open conformation.

Monomer A, which has slightly better density than monomer I, will be used to describe the closed conformation for loop 266−276. In the closed conformation, residues 270−273 make a type I β turn and cover the heme binding site (Figure 2A). Two hydrogen bonds are donated to a propionate side chain of the prosthetic group by the backbone amide groups of Gly272 and Thr274 (Figure 2C). In addition to interactions with the heme group, there are a significant number of hydrogen-bonding interactions between protein atoms found in the closed conformation. Four hydrogen bonds are formed within the loop: between Gly273 (NH) and Gly270 (CO), between Gly270 (NH) and Thr274 (CO), and two hydrogen bonds between the Arg269 guanidinium group and Gly273 (CO). In addition, a water molecule bridges Val277 (NH) and the carbonyl groups of Gly266 and Lys268, forming three hydrogen bonds. Two hydrogen bonds are formed between loop 266−276 and the other glycine-rich loop (138−142) and helix 3_{10}-1: between the side chain of Thr271 and the carbonyl or Ser140, and between the guanidinium group of Arg134 and the carbonyl group of Arg269. Arg269 also makes an electrostatic interaction with the Glu275 side chain; however, the density for the Glu275 side chain is poor, which suggests possible disorder.

In the open form of loop 266−276, the heme group is partially exposed to the solvent, although the clear density

Table 4: Summary of the Glycine-Rich Loop 266−276 from the 16 Monomers[a]

molecule	subunit	conformation	disordered residues
1 (ABCD)	A	closed	
	B		271−274
	C		268−272
	D	open	
2 (EFGH)	E		270−274
	F		270−274
	G		270
	H	open	
3 (IJKL)	I	closed	
	J		269−272
	K		268−273
	L	open	
4 (MNOP)	M		270−274
	N	open	
	O		269−272
	P	open	

[a] The conformation is only indicated for the monomers that have the complete loop built.

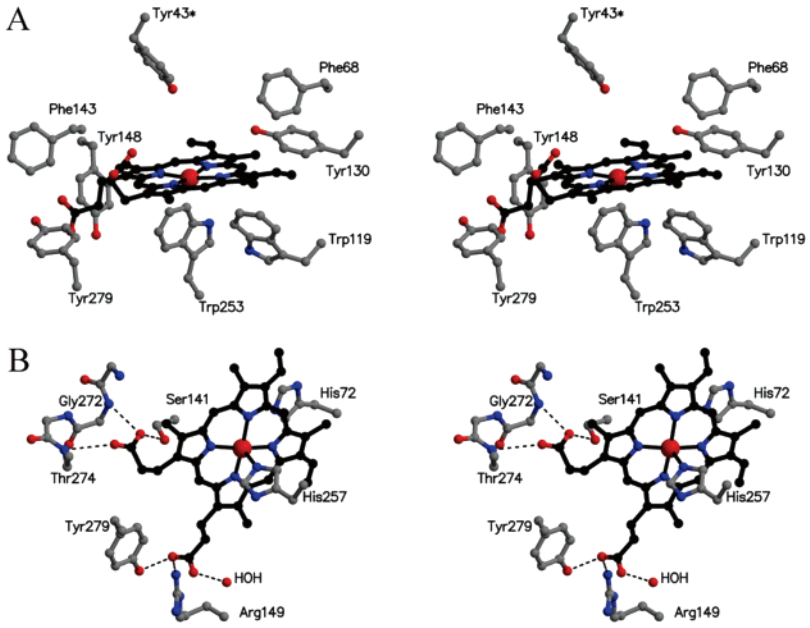


FIGURE 3: The heme binding site. (A) Stacking interactions provided by aromatic residues. (B) Hydrogen-bond interactions.

and relatively low-temperature factors indicate that the prosthetic group is still tightly bound. Monomer H has the best density at the glycine-rich loop and therefore will be used to describe the open form (Figure 2B). Compared to the closed conformation, the open form has fewer direct hydrogen-bonding interactions within the molecule. Only one hydrogen bond, from the amide group of Gly276, is donated by the glycine-rich loop to the propionate group of the heme. The carbonyl group of Gly273 makes two hydrogen bonds with Arg282. There is no direct hydrogen bond formed within the loop; however, water molecules mediate hydrogen bond formation between loop residues. The number of water molecules and their positions vary among monomers. In monomer H, there are five water molecules bridging protein residues from the loop 266−276 forming an intensive hydrogen-bonding network, while in monomer L there is only one ordered water molecule forming a total of three hydrogen bonds.

In addition, the loop in the open conformation forms a hydrophilic interface with a neighboring molecule, related either by translational symmetry or NCS. Monomers D, H, L, and P showed similar hydrophilic interactions between the loop 266−276 and the N-terminal region of an adjacent molecule (Figure 2B), with five or six hydrogen bonds at the interface. Monomer N shows hydrogen bonding interactions at the interface but is different from monomers D, H, L, and P. In monomer N, loop 266−276 forms two hydrogen bonds with the connecting loop between α12 and α13 from an adjacent molecule.

The two monomers adopting the closed conformation, monomer A and I, both have loop 266−276 exposed to the solvent. The five monomers in the ordered open conformation (D, H, L, N, and P) all have favorable crystal packing contacts with an adjacent tetramer. The loops for the other monomers, which either lack interpretable density or have weak density, are all potentially at a different crystal packing contact interface, except monomers E and M for which the loop is solvent exposed. Using the interpretable density for the partially ordered loop, the remaining nine monomers appear more likely to take the open conformation. The observation of two distinct loop conformations and the extent of loop flexibility suggest that the loop likely undergoes a conformational change between the open and the closed forms at some point during the chemical reaction.

Heme Binding Site. TDO has a heme binding site for each monomer that is located at one end of the helical bundle. Six α-helices (α3, α4, α5, α6, α11, and α12), two 3_{10} helices (3_{10}-1 and 3_{10}-2), the two glycine-rich loops, and helix α1 from the adjacent monomer form a mostly hydrophobic cavity. The heme binding geometry is similar between the open and the closed loop 266−276 conformations, except that the hydrogen bonds donated by the loop are different in the two conformations. Monomer A was chosen to describe the heme binding scheme for the closed form.

Eight aromatic residues, Phe68, Trp119, Tyr130, Phe143, Tyr148, Trp253, and Tyr279, and Tyr43* (residues from an adjacent monomer are indicated with an asterisk throughout), line the heme binding cavity and provide stacking interactions with the heme cofactor (Figure 3A). In addition, the side chains of Met79, Leu122, Met125, Val261, Val264, Ile265, and Leu280 contribute to the hydrophobicity of the heme-binding cavity. The two propionate groups of the heme are stabilized by hydrogen-bonding interactions (Figure 3B). The Ser141 side chain and the backbone amide groups of Gly272 and Thr274 donate three hydrogen bonds to one of the propionate groups. The side chains of Arg149 and Tyr279, and a water molecule form three hydrogen bonds to the other propionate group.

His257, which is absolutely conserved in all TDO sequences, is nearly perpendicular to the plane of the heme and coordinates to the iron atom with the Nϵ atom (Figures 2A and 3B). This histidine residue is the only endogenous axial ligand coordinated to the heme, consistent with spectroscopic studies (*33*, *34*). His257 is surrounded by mostly hydrophobic interactions that are provided by Trp119, Trp253, Val261, and Leu280. A water molecule bridges the Nδ of His257 and Arg254 (CO). The other axial position is free in the structure and available for O_2 coordination. The

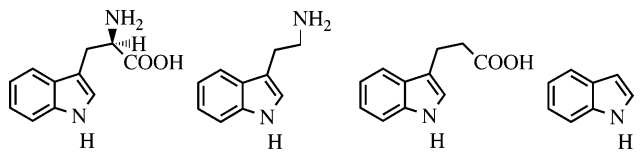
Table 5: The Energies and Torsion Angles from the Modeling

putative intermediate	energy (kJ/mol)	torsion angle
7	−171.03	−173.6
8	−228.06	163.0
9	−172.15	13.8
10	−202.58	−18.9
8 - no torsional constraints	−234.82	90.6
7 - no torsional constraints	−221.39	−51.0

side chain of His72, which is highly conserved, protrudes into a small cavity at the opposite side of the heme relative to His257 and provides a stacking interaction with the heme (Figure 2A). The size of the cavity is appropriate to accommodate the substrates. The side chain of the fully conserved tyrosine residue, Tyr43* from α1 of the adjacent monomer, also inserts into the cavity. This tyrosine residue, which helps to form the heme binding site, is the only residue from the adjacent subunit within a 5 Å radius of the heme. The hydroxyl group of Tyr43* forms a hydrogen bond with Glu76, which in turn hydrogen bonds to Ser145 (OH) and Gln147 Nϵ atom (Figure 2D).

Modeling of the Hydroperoxy Intermediate. Although the substrate L-Trp was present in the crystallization condition, there was no corresponding electron density in any of the 16 subunits. Because attempts to obtain experimental data for a substrate or product complex have so far been unsuccessful, we prepared a model of the hydroperoxy intermediate in the predicted active site (Table 5). The global energy minimum structure for the hydroperoxy intermediate showed two residues with significant conformational change: Tyr43* and His72. The rest of the heme binding site residues are essentially the same as the unliganded TDO structure. The tryptophan ring of the intermediate stacks perpendicularly to the side chain of Phe68. While the minimum energy structure corresponded to a Cβ−Cγ torsion angle in between anti periplanar and syn coplanar, the constrained simulations showed a clear preference for the S-anti periplanar structure (**8**) over structures **7**, **9**, and **10**. In this model, the L-Trp substrate forms good hydrogen bonds between its carboxylate group and Tyr130, Arg134, and the amide group and side chain of Thr271, and hydrogen bonds between its amino group and Thr271 and a propionate of the heme group (Figure 4).

Activities of Mutants and Substrate Analogues. Several mutants and several substrate analogues were tested for activity (Table 6). D-Trp (**11**), tryptamine (**12**), indole propionic acid (**13**), and indole (**14**) were not substrates for TDO. L-Trp was not a substrate for the F68A, R134A, and T271A mutants; however, the Y130F mutants showed a 15-fold increase in activity compared to wild type TDO using L-Trp as the substrate.



11 **12** **13** **14**

DISCUSSION

Comparison with Other TDOs. A BLAST (*35*) search of the sequence database starting with *R. metallidurans* TDO revealed 79 amino acid sequences after removing duplicates and partial sequences. Multiple sequence alignment using CLUSTALW (*36*) showed that the sequences fall into three subgroups. Group 1 contains all 26 eukaryotic sequences and one bacterial sequence (*Bdellovibrio bacteriovorus*). Group 2 is formed by 46 bacterial sequences, including *R. metallidurans* and *X. campestris*. Group 3 has six members, *Cytophaga hutchinsonii*, *Polaribacter filamentus*, *Ralstonia solanacearum*, *Streptomyces filamentosus*, *Nostoc punctiforme*, and *Gloeobacter violaceus*, which have relatively high sequence divergence compared to the first two groups. Members from different groups share approximately 20−30% of identities. The three subgroups of TDOs appear to have higher homology in the N-terminal halves of the sequences (approximately 150 amino acids). The C-terminal halves of the sequences show low intergroup homology. While sequences are more conserved within group 1 and group 2, group 3 has low similarities within the group.

There are 10 absolutely conserved amino acid residues among all 79 sequences: Tyr43, Phe68, Glu76, Arg107, Arg134, Ser141, Gly142, Ser145, His257, and Gly273. All the conserved residues are clustered around the heme binding site except Arg107, which is located at the subunit interface approximately 10 Å away from the heme and may be structurally important rather than catalytically important. In the *R. metallidurans* TDO structure, Phe68, Ser141, Gly142, Gly176, and His257 are involved in heme binding. Phe68 stacks with the heme at its hydrophobic face. Ser141 and Gly142 are on one of the glycine-rich loops wrapping around the heme group, and Ser141 hydrogen bonds to the propionate group of the heme. Gly273 is part of the other glycine-rich loop. His257 is the endogenous axial ligand coordinated to the heme iron atom. Tyr43*, Glu76, and Ser145 appear to only provide stacking and van der Waals interactions to the heme group; however, these three residues also form an interesting chain of hydrogen bonds (Figure 2D). Arg134 is relatively far from the heme binding site, approximately 7 Å away from the heme group, and does not provide direct interactions. However, the side chain of Arg134 has a moderately high-temperature factor, which suggests conformational flexibility. More interestingly, the Arg134 side chain shields from the solvent a small cavity on one side of the heme plane (Figure 2A). In addition to Arg134, the cavity is formed by the heme group, the two glycine-rich loops, Tyr43*, Phe68, His72, and Tyr130. His72 and Tyr130 are both highly conserved. His72 is replaced by a threonine residue in three group 2 bacterial species (*Thermobifida fusca*, *Streptomyces coelicolor*, and *Streptomyces avermitilis*). Tyr130 is replaced by a phenylalanine residue in approximately half of the TDO sequences. The high degree of amino acid conservation around this cavity is consistent with its role as the active site. On the basis of modeling studies, Tyr43*, Phe68, His72, and Arg134 are possibly involved in tryptophan binding and/or catalysis.

Structural Comparison of TDO and IDO. IDO catalyzes the oxidative ring opening of L-Trp but also accepts other substrates while TDO does not. The structure of human TDO was recently reported by Shiro et al. (*12*, *13*). An initial structural homology search for TDO using the DALI server (*37*) revealed IDO at a low similarity score ($Z = 4.8$). A further pairwise comparison of TDO with only the large domain of IDO showed a Z score of 15.3 and a root mean

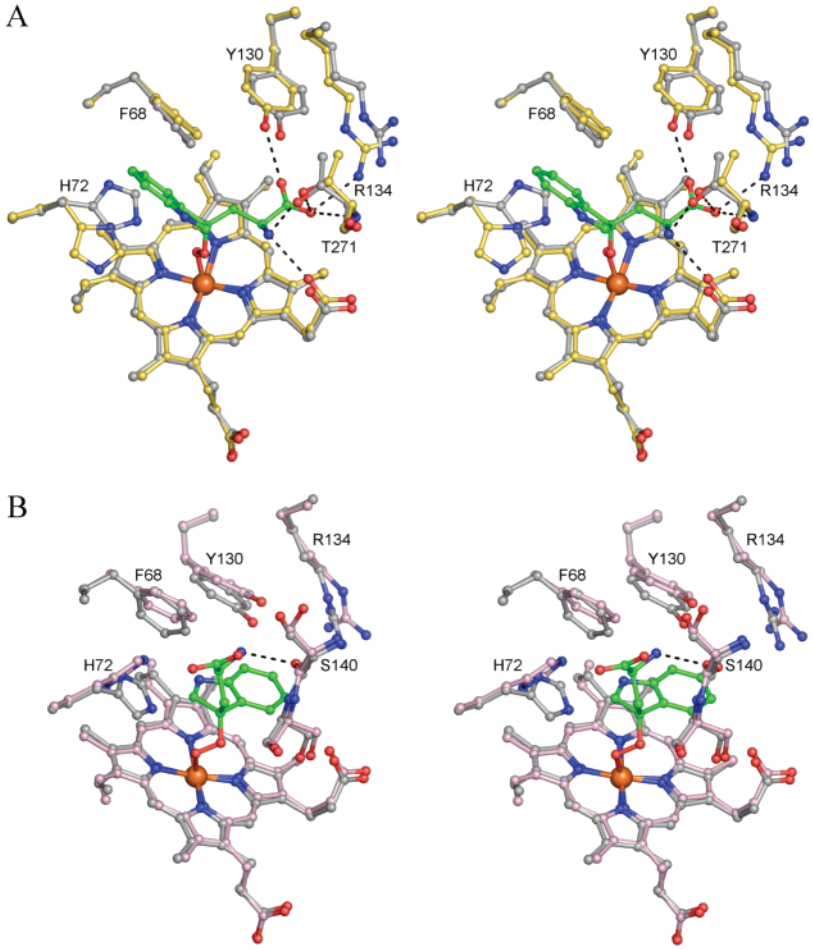


FIGURE 4: Models for the binding of the putative tryptophan hydroperoxide at the active site of TDO. (A) Binding of tryptophan hydroperoxide in the anti periplanar conformation (**8**). (B) Binding of tryptophan hydroperoxide in the syn coplanar conformation (**10**). In both cases, the configuration at the carbon attached to the hydroperoxide is S. The structure of TDO (substrate free) is superimposed and colored in gray for the carbon atoms.

Table 6: TDO Mutants and Substrate Analogues

mutant	substrate	k_{cat}/K_m ($\mu M^{-1} s^{-1}$)	relative activity
native	L-tryptophan (**1**)	3.06×10^{-5}	1
native	D-tryptophan (**11**)	inactive	0
native	tryptamine (**12**)	inactive	0
native	indole propionic acid (**13**)	inactive	0
native	indole (**14**)	inactive	0
F68A	L-tryptophan (**1**)	inactive	0
Y130F	L-tryptophan (**1**)	45.80×10^{-5}	15
T271A	L-tryptophan (**1**)	inactive	0
R134A	L-tryptophan (**1**)	inactive	0

squared deviation of 2.4 Å for 186 aligned residues, even though the sequence identity is only 15% (*38*).

While TDOs function as homotetramers, IDO has been reported as a monomeric protein (*10*) with molecular weight around 45 kDa. Interestingly, the structures of human IDO (PDB codes: 2D0T and 2D0U) each have two monomers in the asymmetric unit related by two-fold NCS, which appear to form a disulfide bond at the monomer/monomer interface (*13*). However, the interface is mostly lined with hydrophilic residues and does not have extensive interactions between the two monomers, which suggests that the disulfide bond is more likely an artifact from the crystal packing. In addition, the monomeric IDO contains a complete active site, and the structure is folded into two distinct domains: the small N-terminal domain comprised of about 155 residues and the large C-terminal domain of 248 residues.

The structure of TDO and the large domain of IDO utilize the same fold with 10 common helices, including α1* and the three long α-helices α3, α4, and α11 (Figure 5A). The heme binding sites also superimpose along with the secondary structural elements. The main difference is that the IDO structure utilizes its small domain to cover the top of the heme binding site, while TDO uses the N-terminal region, especially helix α1* from the adjacent subunit. Human IDO only has one glycine-rich loop (361 GGSAG 365), which is structurally equivalent to residues 139−143 in the TDO structure. IDO residues 360−380 are missing in the structure; however, this region is near the other glycine-rich loop (266−276) in the TDO structure and therefore is potentially the structural equivalent and could also function to close off the active site. IDO has 10 extra residues in this region, which could be responsible for the difference in substrate specificity between TDO and IDO.

Despite the overall low sequence identity, the heme binding environment is mostly conserved between two structures (Figure 5B). The absolutely conserved residues in the TDO substrate binding site are also present in human IDO with the exception of Gly142, which is replaced by an alanine residue. His72 has a structural equivalent of the serine residue in IDO. Tyr130 is replaced by a phenylalanine residue.

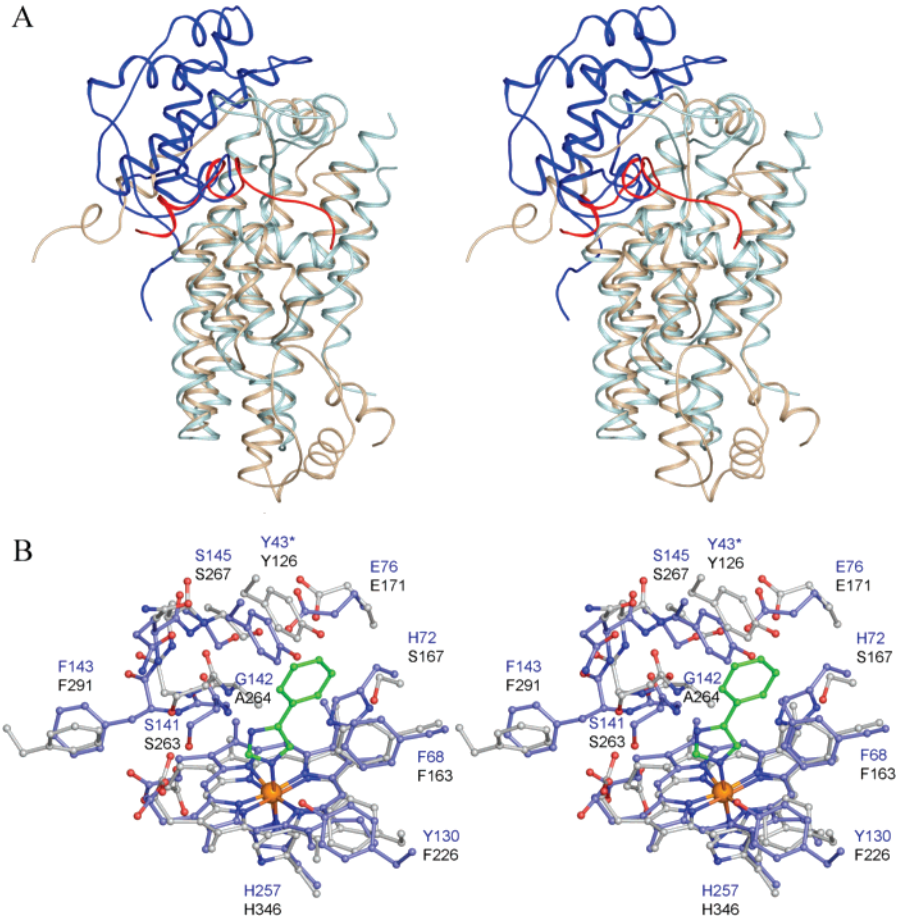


FIGURE 5: TDO and IDO superimposed. (A) The structures of TDO and IDO are shown in thin ribbon diagram. The monomer A of TDO is in wheat color and the N-terminal region including α1 and α2 from monomer B is in red. The small and large domains of IDO are colored in dark blue and light blue, respectively. (B) The heme binding sites are shown in ball and stick with the carbon atoms from TDO in blue and IDO in gray. The ligand 4-phenylimidazole in the structure of human IDO is highlighted in green for the carbon atoms. Structurally equivalent residues are labeled in blue for TDO and in gray for IDO. Conserved residue Arg134 is also present in human IDO (Arg231) but is not included in the figure.

Mechanistic Implications. The mechanism by which TDO catalyzes the oxidative ring opening of the indole ring of tryptophan is still poorly understood. This reaction is significantly different from the well-studied oxidative aromatic ring opening reaction catalyzed by the non-heme dioxygenases. For the latter enzymes, the substrate and molecular oxygen are coordinated to the active site iron, and this coordination is a key part of the catalytic mechanism (*39*). Such coordination is not possible for TDO because the heme has only one vacant coordination site, which is occupied by the oxygen during the reaction. The tryptophan is therefore not coordinated to the heme iron in TDO.

Two mechanistic proposals for TDO are outlined in Scheme 2 (*40*). For both mechanisms, the heme iron catalyzes the formation of the tryptophan hydroperoxide **3** most likely by a single electron-transfer mechanism. In the Criegee mechanism, this hydroperoxide rearranges to give **4** which then collapses to **5**. Ring opening of hemiacetal **5** gives formylkynurenine **2**. In the second mechanism, the hydroperoxide **3** is converted to the dioxetane **6**, which undergoes a retro 2 + 2 cycloaddition to give **2**. This mechanism is considered unlikely due to the high strain in **6**, but it has not been rigorously excluded. An alternative mechanism involving water addition to the C2 carbon of **3** can be eliminated because both oxygen atoms from molecular oxygen are incorporated into the product (*41*).

The Criegee rearrangement (**3** to **4**) requires an anti periplanar arrangement of the O−O bond and the C2−C3 bond of the indole (i.e., **7** or **8**), while the dioxetane mechanism requires a syn coplanar configuration (i.e., **9** or **10**). These four putative intermediates have been modeled into the active site. The resulting energies and torsion angles are given in Table 5. The modeling suggests that the binding of the R-hydroperoxides (**7** and **9**) is less favorable than the binding of the S-hydroperoxides (**8** and **10**) and that binding of the anti periplanar conformation (**8**) is favored over binding of the syn coplanar conformation (**10**). While these results provide some support for the Criegee mechanism, this support is tentative: when the torsional angle constraint is removed, both the R and the S hydroperoxides bind tightly with the oxygen located halfway between the syn and the anti conformations.

The structures generated by modeling the S-hydroperoxides (**8** and **10**) into the active site are shown in Figure 4. For the anti hydroperoxide (**8**) the carboxylate is anchored at the active site by hydrogen-bonding interactions with Arg134, Tyr130, and the hydroxyl and amide NH of Thr271. The tryptophan amino group is hydrogen bonded to the hydroxyl of Thr271 and to the heme propionic acid side chain. The indole ring is stacked against Phe68. Overall, there is good structural complementarity between the intermediate and the active site functionality with only one possible

exception: a basic residue for the deprotonation/reprotonation of the indole nitrogen is not apparent.

For the syn hydroperoxide (**10**), there are no interactions between the carboxylate and the enzyme, the tryptophan amino group is hydrogen bonded to the carbonyl group of Ser140 and the indole is stacked over the heme. A basic residue for the deprotonation/reprotonation of the indole nitrogen is also not apparent in this model. Overall, there is poor structural complementarity between this intermediate and the active site functionality.

The oxidation of substrate analogues, as well as the catalytic properties of active site mutants, were examined to experimentally differentiate between the syn and anti conformations of the bound (*S*)-tryptophan hydroperoxide. The results are shown in Table 6.

The model for the binding of the anti hydroperoxide (**8**) predicts specific strong interactions between the enzyme and the amino and carboxylate groups of the substrate as well as high selectivity for the L-isomer of tryptophan over the D-isomer. In contrast, the model for the binding of the syn hydroperoxide (**10**) predicts weak or no interactions between the enzyme and the amino and carboxylate groups of the substrate as well as low selectivity for the L-isomer of tryptophan. This suggests that that D-tryptophan (**11**), tryptamine (**12**), indole propionic acid (**13**), and indole (**14**) will be good substrates for the enzyme if the oxidation proceeds via the dioxetane mechanism and poor substrates if the oxidation proceeds via the Criegee mechanism. Our observation that none of these analogues are substrates for TDO (Table 6) supports the model for the binding of the (*S*)-tryptophan hydroperoxide in the anti periplanar conformation and suggests that the oxidation occurs by the Criegee mechanism. The results of active site mutagenesis experiments are generally in agreement with this analysis. In the model for the binding of the anti hydroperoxide **8**, Phe68 forms the indole-binding pocket, Arg134 forms hydrogen bonds to the substrate carboxylate group and the hydroxyl of Thr271 forms hydrogen bonds to the substrate amino and carboxylate groups. As predicted, the F68A, R134A, and T271A mutants are inactive. The model suggests that the hydroxyl of Tyr130 also forms a hydrogen bond to the substrate carboxylate. The mutagenesis experiments do not support this interaction because the Y130F mutant shows a 15-fold increase in catalytic activity rather than the expected decrease. All mutants bound the heme cofactor suggesting that loss of activity was not due to protein misfolding. The mutants were also analyzed using the model for the binding of the syn hydroperoxide. The large loss of catalytic activity observed for the F68A, R134A, and T271A mutants is not consistent with this model.

In conclusion, the modeling of the tryptophan hydroperoxide into the active site of the enzyme coupled with substrate analogue and active site mutagenesis studies are consistent with a Criegee mechanism for the indole oxidative ring opening reaction. Additional biochemical experiments are in progress to further elucidate the mechanistic details of this important oxidation.

ACKNOWLEDGMENT

We thank Cynthia Kinsland for the preparation of the mutants, Leslie Kinsland for assistance in preparation of this manuscript, and the staff of beamline X25C at the National Synchrotron Light Source for assistance with data collection.

REFERENCES

1. Tanaka, T., and Knox, W. E. (1959) The nature and mechanism of the tryptophan pyrrolase (peroxidase-oxidase) reaction of *Pseudomonas* and of rat liver, *J. Biol. Chem. 234*, 1162−1170.
2. Kotake, Y., and Masayama, I. (1936) The intermediary metabolism of tryptophan. XVIII. The mechanism of formation of kynurenine from tryptophan, *Z. Physiol. Chem. 243*, 237−244.
3. Takikawa, O. (2005) Biochemical and medical aspects of the indoleamine 2,3-dioxygenase-initiated L-tryptophan metabolism, *Biochem. Biophys. Res. Commun. 338*, 12−19.
4. Ishiguro, I., Naito, J., Saito, K., and Nagamura, Y. (1993) Skin L-tryptophan-2,3-dioxygenase and rat hair growth, *FEBS Lett. 329*, 178−182.
5. Suzuki, S., Tone, S., Takikawa, O., Kubo, T., Kohno, I., and Minatogawa, Y. (2001) Expression of indoleamine 2,3-dioxygenase and tryptophan 2,3-dioxygenase in early concepti, *Biochem. J. 355*, 425−429.
6. Higuchi, K., and Hayaishi, O. (1967) Enzymic formation of D-kynurenine from D-tryptophan, *Arch Biochem. Biophys. 120*, 397−403.
7. Yamamoto, S., and Hayaishi, O. (1967) Tryptophan pyrrolase of rabbit intestine. D- and L-tryptophan-cleaving enzyme or enzymes, *J. Biol. Chem. 242*, 5260−5266.
8. Hirata, F., and Hayaishi, O. (1972) New degradative routes of 5-hydroxytryptophan and serotonin by intestinal tryptophan 2,3-dioxygenase, *Biochem. Biophys. Res. Commun. 47*, 1112−1119.
9. Hayaishi, O. (1976) Properties and function of indoleamine 2,3-dioxygenase, *J. Biochem. (Tokyo) 79*, 13P−21P.
10. Shimizu, T., Nomiyama, S., Hirata, F., and Hayaishi, O. (1978) Indoleamine 2,3-dioxygenase. Purification and some properties, *J. Biol. Chem. 253*, 4700−4706.
11. Hayaishi, O., Hirata, F., Fujiwara, M., Ohnishi, T., and Nukiwa, T. (1975) Catalytic properties and reaction mechanism of indoleamine 2,3-dioxygenase, *Proc. FEBS Meeting 40*, 131−144.
12. Oda, S., Sugimoto, H., Yoshida, T., and Shiro, Y. (2006) Crystallization and preliminary crystallographic studies of human indoleamine 2,3-dioxygenase, *Acta Crystallogr. F 62*, 221−223.
13. Sugimoto, H., Oda, S., Otsuki, T., Hino, T., Yoshida, T., and Shiro, Y. (2006) Crystal structure of human indoleamine 2,3-dioxygenase: catalytic mechanism of O_2 incorporation by a heme-containing dioxygenase, *Proc. Natl. Acad. Sci. U.S.A. 103*, 2611−2616.
14. Kurnasov, O., Goral, V., Colabroy, K., Gerdes, S., Anantha, S., Osterman, A., and Begley, T. P. (2003) NAD biosynthesis: identification of the tryptophan to quinolinate pathway in bacteria, *Chem. Biol. 10*, 1195−1204.
15. Schwarcz, R. (2004) The kynurenine pathway of tryptophan degradation as a drug target, *Curr. Opin. Pharmacol. 4*, 12−17.
16. Stone, T. W., and Darlington, L. G. (2002) Endogenous kynurenines as targets for drug discovery and development, *Nat. Rev. Drug Discovery 1*, 609−620.
17. Mellor, A. L., and Munn, D. H. (2004) IDO expression by dendritic cells: tolerance and tryptophan catabolism, *Nat. Rev. Immunol. 4*, 762−774.
18. Mellor, A. (2005) Indoleamine 2,3 dioxygenase and regulation of T cell immunity, *Biochem. Biophys. Res. Commun. 338*, 20−24.
19. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
20. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
21. Kleywegt, G. J., and Jones, T. A. (1994) Halloween....mask and bones, in *From First Map to Final Model* (Waller, D., Ed.) pp 59−66, SERC Daresbury Laboratory, Warrington, UK.
22. Collaborative Computational Project-Number 4. (1994) The CCP-4 suite: programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
23. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for the building of protein models in electron

density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
24. Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T., and Still, W. C. (1990) MacroModel - an Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics, *J. Comput. Chem. 11*, 460−467.
25. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., and Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins, *J. Am. Chem. Soc. 106*, 765−784.
26. Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) An all atom force field for simulations of proteins and nucleic acids, *J. Comput. Chem. 7*, 230−252.
27. Ponder, J. W., and Richards, F. M. (1987) An Efficient Newton-like Method for Molecular Mechanics Energy Minimization of Large Molecules, *J. Comput. Chem. 8*, 1016−1024.
28. Giammona, D. A. (1984) pp 151, University of California, Davis, Davis, California.
29. Qiu, D., Shenkin, P. S., Hollinger, F. P., and Still, W. C. (1997) The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate born radii, *J. Phys. Chem. A 101*, 3005−3014.
30. Ausubel, F. M., and Brent, F. (1987) in *Current Protocols in Molecular Biology*, John Wiley and Sons, New York.
31. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Vol. 3, Cold Spring Harbor Laboratory Press, Plainview, New York.
32. Erithacus Software, Erithacus Software, Horley, UK.
33. Uchida, K., Shimizu, T., Makino, R., Sakaguchi, K., Iizuka, T., Ishimura, Y., Nozawa, T., and Hatano, M. (1983) Magnetic and natural circular dichroism of L-tryptophan 2,3-dioxygenases and indoleamine 2,3-dioxygenase. I. Spectra of ferric and ferrous high spin forms, *J. Biol. Chem. 258*, 2519−2525.
34. Uchida, K., Shimizu, T., Makino, R., Sakaguchi, K., Iizuka, T., Ishimura, Y., Nozawa, T., and Hatano, M. (1983) Magnetic and natural circular dichroism of L-tryptophan 2,3-dioxygenases and indoleamine 2,3-dioxygenase. II. Spectra of their ferric cyanide and ferrous carbon monoxide complexes and an oxygenated form, *J. Biol. Chem. 258*, 2526−2533.
35. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
36. 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.
37. Holm, L., and Sander, C. (1998) Touring protein fold space with Dali/FSSP, *Nucleic Acids Res. 26*, 316−319.
38. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrixes, *J. Mol. Biol. 233*, 123−138.
39. Bugg, T. D. H. (2003) Dioxygenase enzymes: catalytic mechanisms and chemical models, *Tetrahedron 59*, 7075−7101.
40. Sono, M., Roach, M. P., Coulter, E. D., and Dawson, J. H. (1996) Heme-containing oxygenases, *Chem. Rev. 96*, 2841−2888.
41. Hayaishi, O., Rothberg, S., Mehler, A. H., and Saito, Y. (1957) Studies on oxygenases; enzymatic formation of kynurenine from tryptophan, *J. Biol. Chem. 229*, 889−896.

BI0620095