Three-Dimensional Structure of the Inosine−Uridine Nucleoside *N*-Ribohydrolase from *Crithidia fasciculata*[†,‡]

Massimo Degano,[§] Deshmukh N. Gopaul, Giovanna Scapin, Vern L. Schramm,* and James C. Sacchettini*

Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461

Received December 19, 1995; Revised Manuscript Received March 13, 1996[⊗]

ABSTRACT: Protozoan parasites rely on the host for purines since they lack a *de novo* synthetic pathway. *Crithidia fasciculata* salvages exogenous inosine primarily through hydrolysis of the N-ribosidic bond using several nucleoside hydrolases. The most abundant nucleoside hydrolase is relatively nonspecific but prefers inosine and uridine as substrates. Here we report the three-dimensional structure of the inosine−uridine nucleoside hydrolase (IU-NH) from *C. fasciculata* determined by X-ray crystallography at a nominal resolution of 2.5 Å. The enzyme has an open (α,β) structure which differs from the classical dinucleotide binding fold. IU-nucleoside hydrolase is composed of a mixed eight-stranded β sheet surrounded by six α helices and a small C-terminal lobe composed of four α helices. Two short antiparallel β strands are involved in intermolecular contacts. The catalytic pocket is located at the C-terminal end of β strands β1 and β4. Four aspartate residues are located at the bottom of the cavity in a geometry which suggests interaction with the ribose moiety of the nucleoside. These groups could provide the catalytically important interactions to the ribosyl hydroxyls and the stabilizing anion for the oxycarbonium-like transition state. Histidine 241, located on the side of the active site cavity, is the proposed proton donor which facilitates purine base departure [Gopaul, D. N., Meyer, S. L., Degano, M., Sacchettini, J. C., & Schramm, V. L. (1996) *Biochemistry 35*, 5963−5970]. The substrate binding site is unlike that from purine nucleoside phosphorylase, phosphoribosyltransferases, or uracil DNA glycosylase and thus represents a novel architecture for general acid−base catalysis. This detailed knowledge of the architecture of the active site, together with the previous transition state analysis [Horenstein, B. A., Parkin, D. W., Estupiñán, B., & Schramm, V. L. (1991) *Biochemistry 30*, 10788−10795], allows analysis of the interactions leading to catalysis and an explanation for the tight-binding inhibitors of the enzyme [Schramm, V. L., Horenstein, B. A., & Kline, P. C. (1994) *J. Biol. Chem. 269*, 18259−18262].

More than a million deaths per year are reported from malaria, trypanosomiasis, and other infections caused by protozoan parasites. Vaccines have proven ineffective against trypanosomes since they can evade the immune response by frequently changing a variant surface glycoprotein (Cross, 1975; Vickerman, 1985). Chemotherapeutic treatment for trypanosomiasis is effective in the early stages of the infection, but currently used drugs also exhibit host toxicity since they are structurally similar to DNA intercalating agents or are arsenic-containing compounds. It is hoped that specific antibiotics can be designed against these organisms by targeting essential enzymes of the novel metabolic pathways of protozoan parasites.

One metabolic feature that distinguishes protozoan parasites from the mammals is the lack of a *de novo* biosynthetic pathway for purines (Steiger & Steiger, 1977; Hammond & Gutteridge, 1984; Pizzi & Taliaferro, 1960). Protozoan parasites rely exclusively on purine salvage from the host for DNA and RNA synthesis. These organisms have developed complex salvage pathways to optimize this process. It has been reported that *Crithidia lucillae* expresses a novel cell-surface 3′-nucleotidase/nuclease (Neubert & Gottlieb, 1990) which is proposed to generate nucleosides which are further metabolized by the parasite. A family of nucleoside and deoxynucleoside hydrolases then converts the ribonucleosides to the respective bases and ribose (Dewey & Kidder, 1973; Miller et al., 1984). The purine bases are salvaged with 5-phosphoribosyl 1-pyrophosphate transferase reactions to form the corresponding nucleotides (Kidder et al., 1978; Tuttle & Krenitsky, 1980). An inosine salvage system also exists in mammals; however, the first reaction is a phosphorolysis catalyzed by purine nucleoside phosphorylase (Krenitsky, 1967). Inosine phosphorylase is absent in *Crithidia*, and the nucleoside hydrolases have not been encountered in mammals (Parkin et al., 1991). This metabolic dichotomy between host and parasite offers an opportunity to develop specific inhibitors against the trypanosomal purine salvage pathway.

Two distinct nucleoside hydrolases (inosine−uridine preferring and guanosine−inosine preferring, respectively) have been purified to homogeneity from *Crithidia fasciculata*, a protozoan parasite of the mosquito which provides a convenient laboratory source of protozoan enzymes, since it does not infect mammals (Parkin et al., 1991; Estupiñán & Schramm, 1994). The two enzymes differ in substrate specificity and in their catalytic mechanisms. The IU-NH[1] enzyme derives most of its catalytic potential by converting the ribosyl to a ribooxycarbonium ion, while the GI-NH

[†] This work was supported by Research Grants GM41496 and GM45859 from the National Institutes of Health.

[‡] The atomic coordinates have been deposited in the Protein Data Bank, Brookhaven National Laboratory, PDB code 1MAS (inosine−uridine nucleoside *N*-ribohydrolase).

* Corresponding authors.

[§] Present address: Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1996.

S0006-2960(95)02999-0 CCC: $12.00

obtains nearly equal amounts of transition state stabilization from leaving-group activation as from ribooxycarbonium ion formation (Mazzella et al., 1996). Enzymes with kinetic properties similar to those of IU-NH have also been demonstrated in *Trypanosoma cruzi* (Miller et al., 1984). The transition state structure for the IU-NH from *C. fasciculata* has been determined by multiple kinetic isotope effects and bond-energy bond-order vibrational analysis (Horenstein et al., 1991). Its features include ribooxycarbonium ion character, with an almost full positive charge on the ribose ring, an elongated N-glycosidic bond, and an enzyme-directed water nucleophile attacking the *re* face of the anomeric carbon of the ribose. Protonation of the purine ring is also hypothesized, since the anionic form of hypoxanthine would be a poor leaving group, and ^{15}N isotope effects are consistent with purine ring protonation. The profiles of V_{max} or V_{max}/K_M vs pH show an unprotonated group with a pK_a of 7.1 and a protonated group with a pK_a of 9.1 required for catalysis but not for the binding of substrate (Parkin & Schramm, 1995).

The IU-NH from *C. fasciculata* has been cloned, and the cDNA has been sequenced and translated into the amino acid sequence. No homology with proteins of known function or with known three-dimensional structure has been found; however, several unidentified open reading frames from a variety of organisms show substantial homology in the region of the protein proposed as the catalytic site. The overexpression of IU-NH in *Escherichia coli* allowed the production of milligram quantities of highly purified protein for crystallographic studies (Gopaul et al., 1996). The enzyme has been crystallized in a form suitable for high-resolution structural studies. Here we report the three-dimensional structure of the IU-NH from *C. fasciculata* to a nominal resolution of 2.5 Å. The structure provides insights into the catalytic mechanism and substrate specificity and offers the possibility of developing novel inhibitors based on the knowledge of the active site architecture.

MATERIALS AND METHODS

Protein Purification and Crystallization. Native IU-NH was purified in milligram quantities from frozen *C. fasciculata* cells following the procedure described by Parkin et al. (1991). Recombinant IU-NH, which became available in the later stage of the crystal structure refinement, was purified from extracts of *E. coli* strain BL21(DE3)pLys S transformed with the pET3d-IUNH plasmid according to the procedure outlined in the accompanying paper (Gopaul et al., 1996). Both native and recombinant IU-NH were crystallized using the vapor diffusion hanging drop method, where 4 μL of protein solution at a concentration of 34 mg/mL in 10 mM phosphate/10 mM glycine buffer at pH = 7.4 was mixed with an equal volume of precipitant solution (mother liquor, ML), consisting of 100 mM potassium citrate, 8% poly(ethylene glycol) (Fluka) with an average molecular weight of 4000, and 0.05% sodium azide, placed on a silanized coverslip, inverted and sealed over a well containing 700 μL of ML, and incubated at 22 °C. Crystals of both the native and recombinant IU-NH grew overnight; the native enzyme typically resulted in crystals of maximum dimensions of 0.3 × 0.4 × 0.05 mm while the recombinant enzyme usually gave larger and thicker crystals. The crystals from both the native and the recombinant enzymes belong to the orthorombic crystal system, with unit cell dimensions a = 63.5 Å, b = 131.9 Å, c = 90.1 Å, and $\alpha = \beta = \gamma = 90°$. After analysis of pseudoprecession photographs, using the PHASES package (Furey & Swaminathan, 1990), the space group was assessed to be $P2_12_12$. The crystals contain two subunits of the IU-NH tetramer in the asymmetric unit with a solvent content of 56%. A high degree of nonisomorphism has been found in IU-NH crystals grown in identical conditions, with the c axis ranging from 85.7 to 93.4 Å. A native data set and the heavy atom derivative data sets were collected from crystals of the enzyme from *C. fasciculata* and were used in the initial stages of the structural determination. All the crystals used have the same unit cell parameters as the native data set reported above. A native diffraction data set collected from a crystal of the recombinant enzyme was used for the first electron density map calculation and in all subsequent stages of the structure determination and refinement.

Table 1: Data Collection and Phase Calculation Statistics

derivative	sites	R_{Cullis}[a]	R_{Kraut}[b]	F_H/ϵ[c]	resolution (Å)	R_{sym}[d]	R_{iso}[e]
native					2.5	0.075	
$SmAc_3$ (iso)	4	0.573	0.095	2.04	3.0	0.081	0.09
$SmAc_3$ (ano)	4		0.234	3.15	3.2	0.081	
K_2PtCl_4	4	0.671	0.199	0.90	3.5	0.145	0.21
$NdAc_3$	2	0.627	0.134	1.35	3.5	0.114	0.17

Mean Figure of Merit = 0.638 for 11 966 reflections to 3.2 Å

[a] $R_{Cullis} = \sum||F_{PH}(hkl)| \pm |F_P(hkl)|| - |F_H(hkl)||/\sum||F_{PH}(hkl)| \pm |F_P(hkl)||$, where F_{PH}, F_P, and F_H are the derivative, native, and calculated heavy atom structure factors, respectively. The sum is extended over all centric reflections. [b] $R_{Kraut} = \sum(|F_{PH}(hkl)| - |F_H(hkl)|)|/\sum|F_P(hkl)|$. The sum is extended over all acentric reflections. [c] F_H/ϵ = phasing power, defined as the ratio of the heavy atom structure factor divided by the lack-of-closure error. [d] $R_{sym} = \sum|(I_i(hkl) - \langle I_i(hkl)\rangle)|/\sum I_i(hkl)$, where $I_i(hkl)$ are the intensities of multiple measurements and $\langle I_i(hkl)\rangle$ is the average of the measured intensities for the ith reflection. [e] $R_{iso} = \sum|(F_{PH}(hkl) - F_P(hkl))|/\sum F_P(hkl)$; the summation is extended over all reflections common between the two data sets.

Data Collection and Processing. X-ray diffraction data sets were collected at room temperature from single crystals using a Siemens X-1000 area detector coupled with a Rigaku rotating anode X-ray generator operating at 80 mA and 55 kV, using the Cu Kα wavelength (1.542 Å). Indexing, integration, reduction, and scaling of the data were performed using the program XENGEN 2.1 (Howard et al., 1987). Data collection statistics for native and derivative crystals are presented in Table 1. The native data set used for the structure determination and refinement has an overall 94% completeness to 2.5 Å resolution and an R-factor on symmetry-equivalent reflections (R_{sym}) of 0.08. The data set is 84% complete between 2.9 and 2.5 Å resolution, with an R_{sym} of 0.18 and an $\langle I/\sigma(I)\rangle$ of 3.6 in this resolution bin.

Heavy Atom Derivatives. More than 40 compounds were screened in order to obtain isomorphous heavy atom derivatives suitable for phasing. Since even the native crystals were suffering from a high degree of nonisomorphism, as

[1] Abbreviations: IU-NH, inosine−uridine-preferring nucleoside hydrolase from *Crithidia fasciculata*; GI-NH, guanosine−inosine-preferring nucleoside hydrolase from *C. fasciculata*; PNP, purine nucleoside phosphorylase; OPRTase, orotate phosphoribosyltransferase; HG-PRTase, hypoxanthine−guanine phosphoribosyltransferase; PRTase, phosphoribosyl transferase; R-factor, $\sum(|F_o(hkl)| - |F_c(hkl)|)|/\sum|F_o(hkl)|$; NCS, noncrystallographic symmetry; MIR, multiple isomorphous replacement.

previously described, the derivative crystals were used only if the unit cell parameters were within 1% of the values found for the native enzyme. The first derivative was obtained after soaking a crystal in ML containing 3 mM samarium acetate for 12 h. Two major sites were found using isomorphous difference Patterson maps. These two sites were used to phase a self-difference Fourier map using the program PHASES, which revealed two minor sites. The four heavy atom sites were confirmed using the program VERIFY (S. L. Roderick, Albert Einstein College of Medicine) which searches for verification of the heavy atom self- and cross-vectors. A second derivative was produced by soaking overnight a crystal in ML containing 1 mM neodymium acetate. Isomorphous difference Patterson maps showed the location of two sites, both coincidental with the major samarium sites described above. A third isomorphous heavy atom derivative was obtained by soaking a crystal in ML containing 5 mM potassium tetrachloroplatinate for 6 days. Four platinum sites were identified in difference Fourier maps calculated with phases from the samarium derivative.

MIR Phasing. The heavy atom parameters of the three derivatives described above were refined using a maximum likelihood function as implemented in the program PHASIT, included in the PHASES suite. Phases were calculated to 3.2 Å using the three derivatives and including the anomalous signal from the samarium. Phases of acentric reflections were included only if at least two derivatives contributed to the phase calculation. The mean figure of merit for the 11 966 phased reflections to 2.5 Å was 0.638. Solvent flattening (Wang, 1985) was used to further improve the phases using the standard conditions implemented in the PHASES suite, using a solvent content of 50%. The MIR phases were extended to all the measured reflections to 3.0 Å resolution in 20 additional cycles of solvent leveling. The final mean figure of merit for the 12 344 reflections was 0.82. The statistics of data collection and phase refinement are summarized in Table 1.

Noncrystallographic Symmetry Determination. The position of the non-crystallographic symmetry axis was determined by searching the asymmetric unit for the highest correlation of electron density using the program DENSCOR (written by M. Degano and S. L. Roderick, Albert Einstein College of Medicine). The search produced a clear single solution, corresponding to a 2-fold axis of symmetry perpendicular to the crystallographic *c* axis and intersecting it at $z = 58.4$ Å. The noncrystallographic symmetry was confirmed by solving the self-rotation function, calculated using the program GLRF (Tong & Rossmann, 1990).

Model Building and Electron Density Improvement. Only the electron density associated with one of the two molecules in the asymmetric unit was readily interpretable, in both unaveraged and globally averaged electron density maps. A polyalanine chain containing 180 residues was manually fit to this electron density map using the program TOM, a derivative of FRODO (Jones, 1985), running on a Silicon Graphics Crimson computer. In order to generate the second molecule in the asymmetric unit, this partial model was rotated and translated according to the NCS operator. The correctness of the operation was confirmed by visual inspection of the MIR map contoured at 0.7σ, showing recognizable density for the secondary structure elements. The rotated model was used as an aid for tracing the poorly defined region of the second molecule. Phases were then calculated using the partial model of the two protein subunits, combined with the MIR phases (Read, 1986) and subjected to 60 cycles of NCS averaging with DEMON (Vellieux, 1993). The mask used in this procedure was based on the partial protein structure. The improved electron density allowed the tracing of 290 of the 315 amino acid residues as polyalanine in each of the subunits. The phase improvement during the tracing of the model was monitored by the increase of the signal associated with the samarium sites in a Fourier difference map calculated with amplitutdes $|F_{Sm}| - |F_P|$ and model phases, using the PHASES package.

Structure Refinement. This partial polyalanine structure was subjected to a simulated annealing procedure as implemented in X-PLOR (Brünger et al., 1987; Brünger, 1992), which resulted in a crystallographic *R*-factor of 0.31 for all data with $F > 2\sigma$ in the resolution range between 10.0 and 3.0 Å. An electron density map calculated with combined model and MIR phases permitted the correct amino acid sequence to be introduced. The model was manually adjusted to the improved density, which also allowed the introduction of 10 of the 25 missing residues. A second simulated annealing, followed by 15 cycles of restrained temperature factor refinement, was carried out using data between 10.0 and 2.5 Å, resulting in an *R*-factor of 0.26. At this stage $(2|F_o| - |F_c|, \phi_c)$ and $(|F_o| - |F_c|, \phi_c)$ maps were used for subsequent cycles of manual model adjustment, using the program O (Jones, 1989). These electron density maps permitted the inclusion of all the remaining amino acids in the model. Several cycles of restrained least squares positional and temperature factor refinement, using the program TNT (Tronrud et al., 1988), reduced the *R*-factor to 0.22. Sixty ordered solvent molecules were included in the final model on the basis of the following criteria: (i) spherical electron density at least three times the standard deviation of a $|F_o| - |F_c|$, ϕ_c map; (ii) distance of ~3.0 Å from a hydrogen bond donor or acceptor; (iii) temperature factor after refinement not greater than 70.0 Å^2.

Postrefinement Analysis. The final model was checked extensively for misfitted residues with the aid of simulated annealed omit maps, calculated with the program X-PLOR. The heavy atom sites were confirmed by inspection of Fourier difference maps, calculated with $(|F_{PH}| - |F_P|, \phi_c)$ coefficients, where F_{PH} and F_P are the amplitudes for the heavy atom derivative and the native, respectively, and ϕ_c are the phases calculated from the refined model. These maps were calculated using the PHASES package. The stereochemical quality of the model was judged with the program PROCHECK (Laskowski et al., 1993).

RESULTS

Heavy Atom Binding Sites. The location of the heavy atoms used for phasing of the MIR density map has been confirmed with difference maps, calculated with $|F_{PH}| - |F_P|$ amplitudes and phases from the refined model. The two major samarium binding sites are found in the two proposed active sites (see below) present in the asymmetric unit, coordinated by three aspartate residues. The two neodymium atoms also bind to the same site. One of the two minor samarium binding sites is found at the surface of contact between chain A and one of its symmetry-related molecules. The lanthanide is chelated by the carboxyls of Glu 220 and Asp 223 from chain A and from Glu 27 of the symmetry-

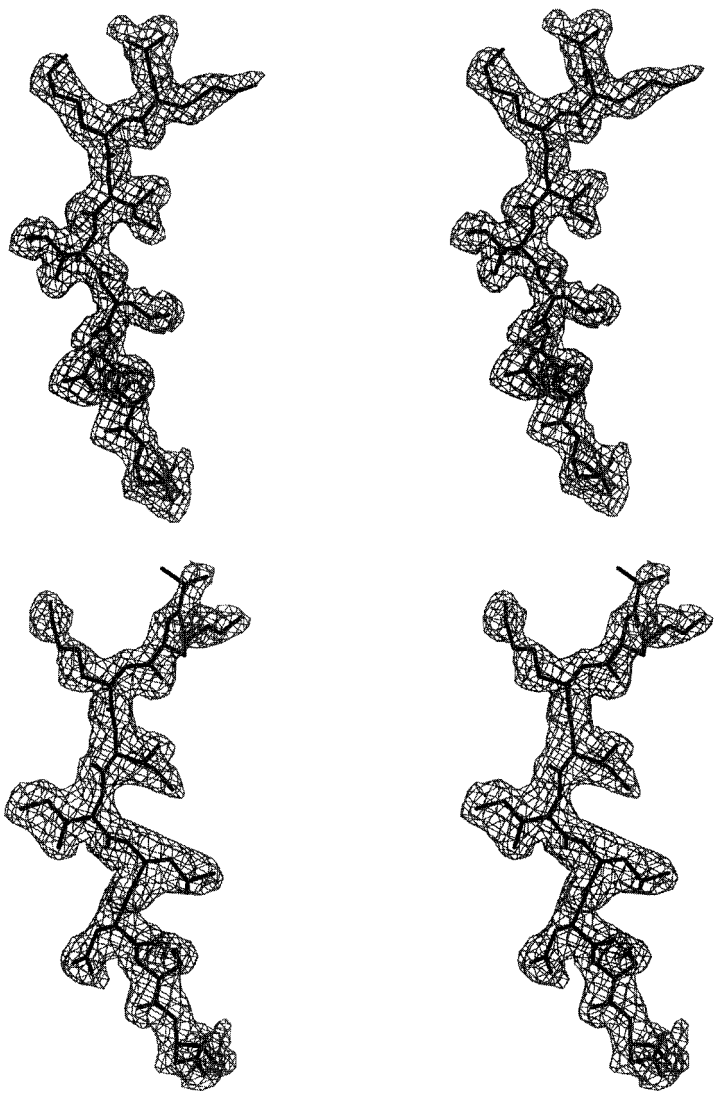

FIGURE 1: Stereo diagram of an example of the final electron density corresponding to strand β1, residues 2−10, of each molecule in the asymmetric unit of the crystal. Residues from subunits A and B are shown in the upper and lower panels, respectively. The map was calculated using coefficients of the type ($|F_o| - |F_c|$, ϕ_c), with phases calculated from the final refined model, after omission of residues 78−83 and 20 cycles of restrained least squares refinement. The electron density is contoured at a level equal to 1 SD of the map.

related polypeptide chain. The same arrangement of residues is not found at the surface of chain B, and this explains why a samarium atom is not found in the same location on the second subunit. The other samarium atom is bound between the carboxyl groups of Glu 102 in chain A and Glu 102 of a molecule related to chain B by crystallographic symmetry.

The platinum atoms bind with different affinities to the two molecules in the asymmetric unit of this crystal. One platinum atom is found bound to the side chain of His 241 in both polypeptide chains in the asymmetric unit; this residue is located in the proposed catalytic pocket. However, the other two binding sites located by analysis of the Patterson maps are bound to the side chains of His 116 in chain A and to the sulfur atoms of Met 218 and Met 222 also in chain A. Inspection of a Fourier difference map between the data set from the crystal soaked in platinum and the native data set calculated with phases from the refined model indicated that two additional platinum atoms are bound to the same residues in chain B, but with lower levels of occupancy.

Quality of the Model. All 628 amino acid residues of the two subunits of the IU-nucleoside hydrolase tetramer found in the asymmetric unit were refined using least squares methods against X-ray diffraction data from a single crystal. The data set used in the refinement (Table 1) contained 89% of all possible reflections in the resolution range 99−2.5 Å with $F > 2\sigma$, and the final model shows a conventional crystallographic *R*-factor of 0.17. The root mean square deviations of the bond lengths and angles from ideal stereochemistry, based on a dictionary obtained from a small molecule database, are 0.016 Å and 1.51°, respectively. The mean temperature factor values for the protein atoms of each IU-NH subunit found in the asymmetric unit are 18 and 38 Å^2, respectively. An example of the electron density at the last stage of refinement for both molecules in the asymmetric unit is presented in Figure 1. Analysis of the refined structure showed that all of the (ϕ, ψ) main chain torsional angles of the IU-NH subunits fall in the allowed regions of a Ramachandran plot (Figure 2). Residues 78 through 83 did not have interpretable electron density associated with them in either molecule in the asymmetric unit and were omitted from the final model. Residual density between residues 78 and 83 is present in an ($|F_o| - |F_c|$, ϕ_c) map contoured at three times its standard deviation; however,

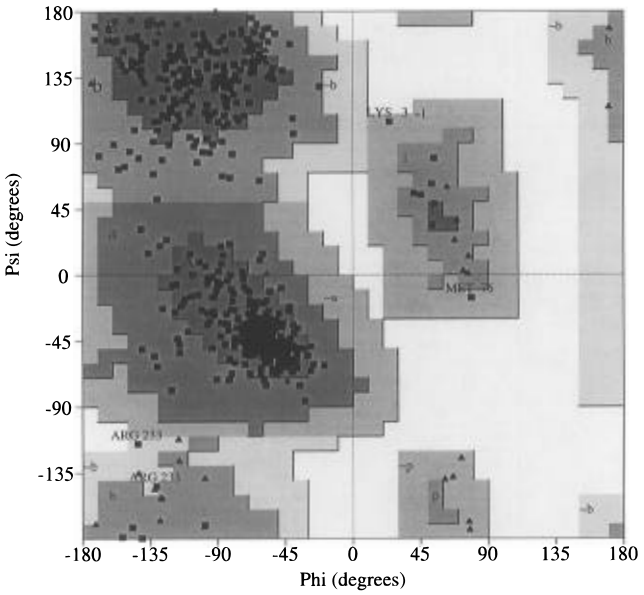

FIGURE 2: Ramachandran plot for all residues except 78−83 in both subunits of the asymmetric unit. No residues fall in prohibited regions. Arg 233 is in a loop extending in a solvent channel and displays some flexibility. Met 76 is the residue before the break in the electron density. Lys 3 is in the N-terminal penultimate amino acid in the Met-processed subunits. These residues fall in marginally allowed regions of the Ramachandran plot. The plot was generated using the program PROCHECK (Laskowski et al., 1993). Glycine residues are shown as triangles. Residues in "most favored", "additionally allowed", and "generously allowed" regions are 86.3%, 13.0%, and 0.8%, respectively.

modeling of the missing residues according to this density does not improve either the *R*-factor or the quality of the ($2|F_o| - |F_c|$, ϕ_c) electron density map after refinement. Moreover, inspection of the resulting ($|F_o| - |F_c|$, ϕ_c) electron density map suggests that at least another conformation of the loop is present in this crystal form. Crystallographic studies at higher resolution are required to be able to model these alternate conformations of this flexible loop.

Structure of IU-Nucleoside Hydrolase. Each subunit of the IU-NH tetramer is folded into a single-domain globular structure (Figure 3) that can be divided into three regions. The first consists of an (α,β) core comprising segments β1−β6, α1−α6, β7, β10, and α10. The second is a lobe consisting of the four α helices α7, α8, α9, and α11. The third region consists of two short segments of β structure, strands β8 and β9, and forms the subunit−subunit interface. A chart showing the overall topology of the enzyme is provided in Figure 4.

The core of the IU-NH subunit contains an eight-stranded central β sheet with six intervening α helices (α1−α6). The first six β strands (β1−β6) and the intervening helices (α1−α6) form the core region of IU-NH and are arranged in a structure that resembles a classical dinucleotide binding or Rossmann fold (Adams et al., 1970). β strands β7 and β10 form an extension to the C-terminal "half-sheet" of the core Rossmann fold. The topology of the central β sheet is (+1*x*, +1*x*, −3*x*, −1*x*, −1*x*, −2, +1) (Richardson, 1981) which has not been reported in other protein structures. β strands β2, β3, β4, β5, β6, and β10 run parallel to strand β1 while strand β7 is antiparallel to all the others. β strands β1−β3 form the N-terminal portion of the core β sheet. These strands are linked by α helices forming the classical right-handed β−α−β motif. A long loop, partly involved in crystal contacts, followed by the third α helix, connects the N-terminal region to the second half of the β sheet. This loop extends into solvent in both polypeptide chains found in the asymmetric unit and is the only region in the protein that is not well-defined. Residues 78−83 have weak or no corresponding electron density, possibly because of conformational flexibility, and have been consequently omitted from the final model. The second half-sheet of the core consists of four parallel β strands and one antiparallel β strand: strands β4 and β5 are joined by an α helix that follows the same right-handed pattern, while between strands β5 and β6 there are two α helices. β strands β7 and β10 complete the core β sheet. An integral part of the central core is helix α10, which runs antiparallel to β strands β1 and β4 and is bordering the proposed catalytic pocket (see Discussion).

Helices α7, α8, α9, and α11 form a lobe on the surface of the protein that wraps one side of the core β sheet. Helix α7 is only eight residues long, 191−197, and is followed by a sharp turn that forces the following helix, α8, to orient nearly perpendicular to α7. The third α helix of the bundle, α9, runs antiparallel to α8, with an interhelix angle of ∼30°. The C-terminal helix, α11, is parallel to α8, interacting via the side chains of the hydrophobic face of this amphipathic α helix. The four α helices of the lobe are not involved in any interactions with other subunits, and their function in the enzyme appears to be purely structural, shielding the hydrophobic residues of the core domain from the solvent.

Two short antiparallel β strands, β8 and β9, extend from the core domain to the interfacial region of the tetramer, defining a plane which is roughly perpendicular to the direction of the last two β strands of the core sheet. The interactions between the two NCS-related subunits are largely mediated by residues from β8 and β9. The loop connecting β5 to α5, residues 156−166, is also involved in interactions between subunits related by crystallographic symmetry, as discussed in the following section.

Quaternary Structure of IU-Nucleoside Hydrolase. The IU-NH elutes from gel filtration chromatography as a homotetramer (Parkin et al., 1991). This arrangement is consistent with the subunit packing in the crystal lattice. The four subunits in the IU-NH molecule are arranged to form a homotetramer of 222 point symmetry (Figure 3). The enzyme crystallized with one of the molecular symmetry axes (*p*,*q*,*r*) coincident with the crystallographic 2-fold of this space group, the *z* axis. The other two molecular axes are perpendicular to this crystallographic symmetry element. The two subunits of the enzyme found in the asymmetric unit of the crystal are related by a 180° rotation about an axis with spherical polar coordinates (Rossmann & Blow, 1962) ϕ = 34.2° and ψ = 90.0°.

The interface buried by the interaction of the two subunits contained within one asymmetric unit is 1045 Å^2. The interactions at this interface, dominated by hydrophobic contacts, are mainly between β9 and its NCS-related counterpart[2] and between the loop joining the C-terminal end of strand β5 to the N-terminal residues of helix α5 and the same region in chain B. Contacts at the van der Waals surface are made by the side chain atoms of residues Val 164 and Leu 270 of chain A with Val 276 of chain B and

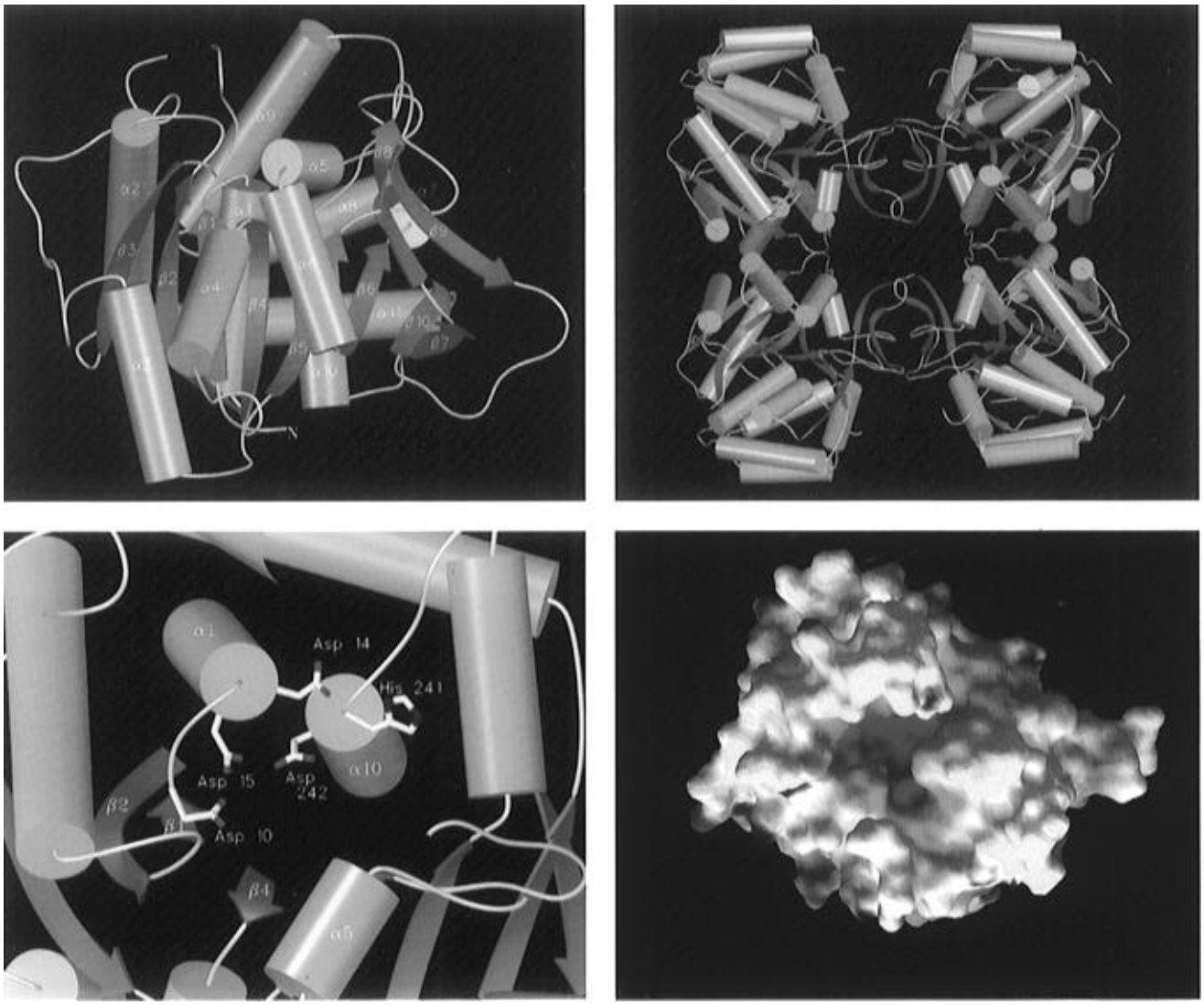

FIGURE 3: The upper left panel shows a front view of the IU-NH monomer with β strands drawn as arrows, α helices as cylinders, and loop regions as lines, labeled as in the text. The upper right panel shows the IU-NH tetramer as generated from the crystallographic symmetry. The *p* and *q* molecular 2-fold axes lie in the plane of the page, horizontal and vertical, respectively. The active sites from opposite subunits are pointing toward the same face of the plane containing the two axes. The lower left panel shows a close-up of the binding cavity with the ionizable residues highlighted. The cluster of aspartate residues is a candidate for interactions with the ribose hydroxyls, while the histidine 241 is a candidate for the general acid in the reaction. The three panels above were generated using the program SETOR (Evans, 1993). The lower right panel shows the molecular surface of the IU-NH monomer, generated with GRASP (Nicholls et al., 1991) and colored according to electrostatic potential (blue = positive, red = negative). The bound metal ion has been removed for this calculation. The proposed catalytic site has a strong negative potential, consistent with the need to stabilize the positively charged transition state of the catalyzed reaction.

by Met 274 with its counterpart from the other molecule. These contacts shield a hydrophobic patch on the surface of the protein from the solvent. Within this patch, hydrogen bonds are formed by the side chains of Asp 263 from chain A with Lys 269 from chain B and by the side chain carboxylate of Glu 265 with the main chain nitrogens of residues Lys 269 and Leu 270 of chain B. The main chain carbonyl oxygen of Thr 267 and the side chain OG1 atom of the same residue in the second molecule form hydrogen-bonding interactions. Additionally, there are hydrogen bonds formed between the main chain carbonyl oxygens of Asn 160 and Ala 161 with the guanidinium group of Arg 283. The imidazole ring of His 157 is bonded with the main chain carbonyl oxygen of the same residue in chain B.

The interface is less extensive between the subunits of the tetramer related by crystallographic symmetry. This interface buries approximately 870 Å^2 of surface area. The interactions are exclusively between hydrophobic residues, mainly between side chain atoms from the loop at the C-terminal end of strand β3. The side chains of Leu 71 and Val 72 are both in contact with the side chain of Ile 178 of a symmetry-related subunit, and there are additional van der Waals contacts between Val 72 and Ile 264. Leu 138 is in contact with Val 106, and Met 134 and the CG1 atom from Thr 267 interact with their counterparts in the crystallographic-related molecule.

The root mean square deviation between the Cαs and all of the atoms in the two subunits in the asymmetric unit is 0.46 and 1.12 Å, respectively. Most of the differences occur

[2] The two identical polypeptide chains found in the asymmetric unit of the crystal are referred to in the paper as chain A and chain B, according to the nomenclature followed in the Protein Data Bank deposition form.

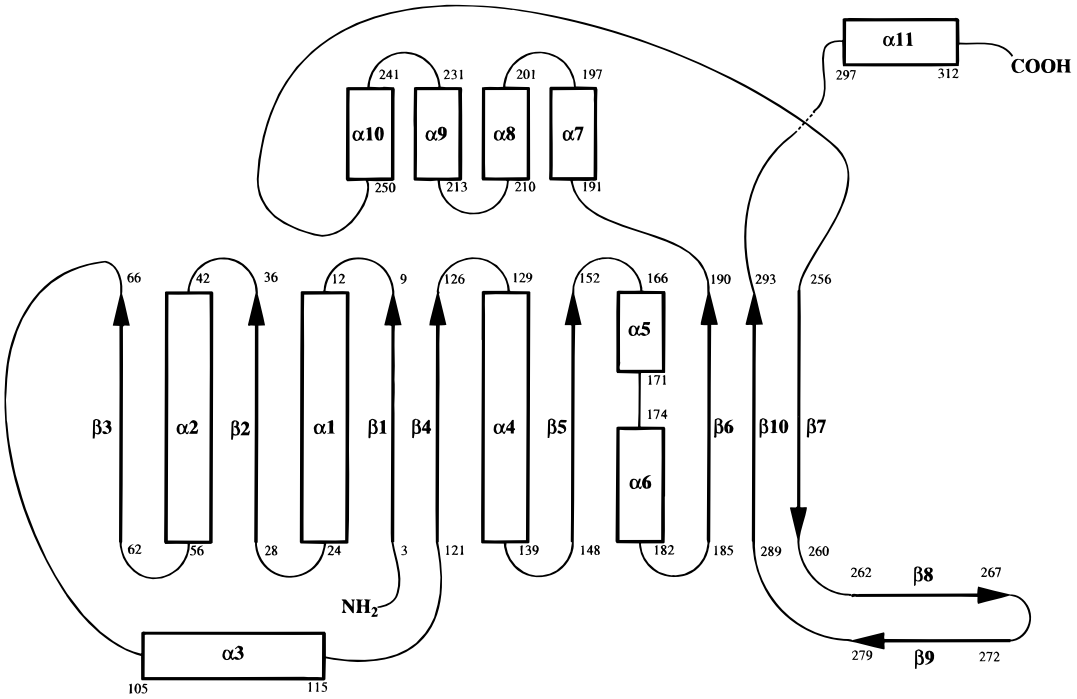

FIGURE 4: Topology diagram of the secondary structure of IU-NH. β strands are depicted as arrows and α helices are drawn as rectangles.

in the loop regions, in particular the long and highly disordered loop between β3 and α3.

In the asymmetric unit of the crystal, one of the two molecules displays a higher degree of disorder, as judged by the weaker electron density surrounding the atoms and the higher temperature factors. The mean temperature factors calculated using all the atoms in each molecule are 18 and 38 Å^2, respectively. The higher mean temperature factors associated with the atoms in one molecule are related to a reduction in the lattice contacts for this subunit, providing increased structural flexibility for the molecule as a whole. This might explain the incidence of nonisomorphism observed in the native crystals. Slight changes in the orientation of the subunits in the asymmetric unit could be achieved by small hinged rotations about the residues at the subunit interface, without significantly changing the packing interactions. A plot of the temperature factors versus residue number for each molecule is provided in Figure 5.

Proposed Active Center. A relatively large cavity is located at the C-terminal end of β1 and β4 (Figure 3). The cavity is roughly cylindrical, with a diameter of 12.5 Å and approximately 10 Å depth, for a volume of about 1227 Å^3. This cavity is found at the region identified as the "topological switchpoint" of the core region of IU-NH. In other enzymes which contain an open (α,β) structure this region is invariably associated with the binding or active site [Adams et al. (1970) and more recent review by Bränden (1980)]. The walls of the cavity are formed by residues from β1, β2, β4, and β5 and also from helices α2 and α9. The side chains of residues Leu 13, Asn 39, Phe 167, Asn 168, Ile 171, Tyr 224, Ile 228, Tyr 229, and His 241 contribute to the internal surface of the cavity. From modeling studies, Phe 167 seems to be in proper orientation to interact with the purine substrate in a base-stacking interaction, as typically observed in proteins which bind nucleotides [e.g., Eads et al. (1994)].

Four aspartate residues, Asp 10, Asp 14, Asp 15, and Asp 242, are located at the bottom of this cavity (Figure 3). The

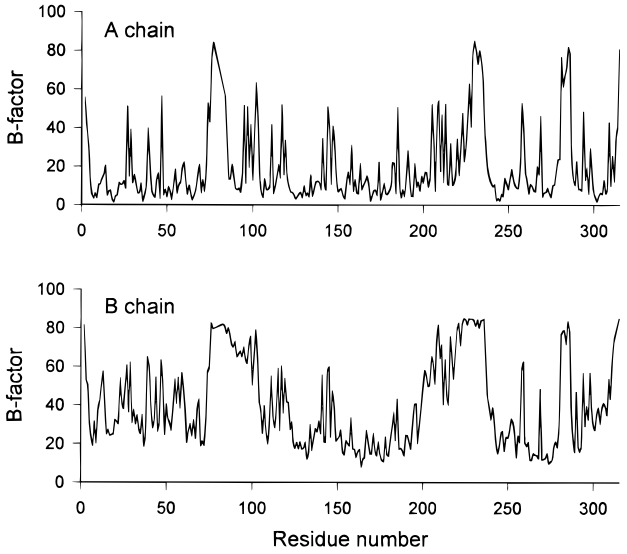

FIGURE 5: Plot of the mean temperature factor of each residue with residue number for the two molecules found in the asymmetric unit. The variation of the temperature factor within one chain follows the same pattern in both molecules. However, the average value of the temperature factors is significantly higher in the B chain.

side chain carboxylate groups of three of the aspartates, Asp 14, Asp 15, and Asp 242, coordinate a sphere of strong density in difference electron density maps calculated with phases from the refined structure. This electron density does not appear to be due to an ordered water, because of its more intense electron density relative to the other bound waters in the structure and its coordination by six ligands, instead of the four hydrogen bonds typically found for an ordered water molecule. Because of the high concentration of potassium ion (0.3 M) in the crystallization mother liquor, we have modeled and refined a bound potassium ion at this location (Figure 6). This ion could possibly be a divalent cation (e.g., calcium); however, but calcium ion frequently involves seven oxygen ligands. The refined temperature factors for the two metal ions in the asymmetric unit are 20

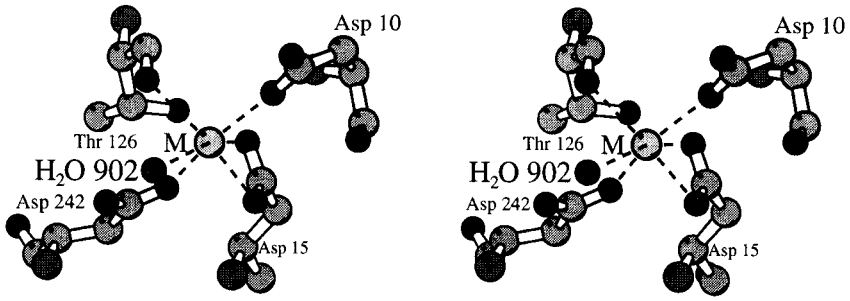

FIGURE 6: Ball-and-stick stereo projection of the residues involved in the binding of the metal ion. The metal−ligand distances are summarized in table 2. This figure has been created using the program MOLSCRIPT (Kraulis, 1991).

Table 2: Interactions between IU-NH and the Metal Ion[a]

ligand	ion−ligand distance (Å)	temperature factor (Å^2)
Asp 242 OD2	2.54	14.5
Asp 15 OD2	2.69	10.3
Asp 15 OD1	2.54	7.0
Asp 10 OD1	2.50	11.7
Thr 126 O	2.71	6.6
H_2O 902	2.43	27.5

[a] The values are for the A subunit. Distances from the B subunit are the same but the temperature factors are greater.

and 35 Å^2, respectively, and correlate well with the temperature factors of the coordinating residues. This ion binding site also corresponds to the location of the two major samarium sites and the two neodymium sites in the crystals soaked with heavy atoms. The coordination of the ion is distorted octahedral, with five protein ligands and an ordered water molecule completing the coordination sphere. The residues involved in coordinating the metal are summarized in Table 2. There is no kinetic evidence for the requirement of added metal ions for IU-NH activity; thus a bound K^+ ion could serve to counterbalance the negative charge present at the bottom of the pocket. An alternative explanation is a tightly bound cation, always associated with the enzyme.

A histidine residue (His 241) at the N-terminal end of helix α10 is located near four aspartate residues and displays weak interactions with Asp 14 (distance NDE−OD1 = 3.29 Å). The arrangement of the carboxylate and histidine residues inside this cavity suggested that this histidine could act as the general acid to protonate the purine ring of the leaving product. Site-directed mutagenesis studies of His241Ala confirmed that this residue is important for inosine hydrolysis (Gopaul et al., 1996). The full catalytic activity of this mutant with *p*-nitrophenyl β-D-riboside is consistent with the hypothesis that His 241 acts as the general acid in the hydrolysis reaction catalyzed by IU-NH (Mazzella et al., 1996).

The proposed active sites from two opposite subunits face the same side of the plane containing the (*p*, *q*) molecular 2-fold axes. The nearest guanidinium nitrogen of Arg 280 from the adjacent subunit is located approximately 15 Å from the opening of the cavity but has sufficient flexibility to reach into the cavity. This residue can be considered a candidate to act as a general acid for leaving-group assistance. A significant movement of the subunits would be necessary to bring the arginine in proper orientation for catalysis. However, as stated above, the His241Ala mutagenesis studies suggest a primary role for the imidazole side chain of this residue.

The molecular surface of the IU-NH subunit, generated with GRASP (Nicholls et al., 1991), colored according to the electrostatic potential, is shown in Figure 3. The proposed catalytic site ion has been removed to illustrate the electrostatic potential of the empty pocket. The proposed catalytic site has a strong negative potential, thus providing a suitable environment to stabilize the positively charged oxycarbonium-like transition state, known for this enzyme (Horenstein & Schramm, 1993a).

DISCUSSION

Comparison of Catalytic Site Structure. The overall fold of IU-NH from *C. fasciculata* resembles the classical dinucleotide binding domain found not only in dehydrogenases but also in enzymes with related functions like the phosphoribosyltransferases (Scapin et al., 1994; Eads et al., 1994; Smith et al., 1994) which bind nucleotidyl monophosphates. IU-nucleoside hydrolase, however, does not display affinity for nucleotides (Parkin et al., 1991). In OPRTase, HGPRTase, and glutamine amido-PRTase most of the interactions between protein and the 5′-phosphate of the bound nucleotides occur at a loop which connects a central β strand of the core to an α helix. This loop is in the second or C-terminal half-sheet of the modified Rossmann fold of the three PRTases, and it is well conserved among most of the PRTases sequenced to date. For example, in the crystal structure of OPRTase with bound orotate monophosphate from *Salmonella typhimurium*, the oxygens of the phosphate group of the nucleotide form six hydrogen bonds with the main chain amide nitrogens of residues Val 126−Thr 131, and two hydrogen bonds with the side chains of Thr 128 and Thr 131. In IU-NH the analogous loop between β5 and α5 is involved in subunit−subunit interactions and displays lower flexibility. Moreover, a negatively charged phosphate could not be accommodated in the proposed substrate binding site of IU-NH since all ionizable groups in the ribosyl region are carboxylates.

Despite the similar chemistries of their catalytic reactions, IU-NH and PNP show no similarity in three-dimensional structure. PNP is also an open (α,β) protein (Ealick et al., 1990), but the topology of the core β sheet is (+3*x*, −1, −1, −2*x*, −2, −1*x*, +2). There is also no significant similarity between the residues in the active sites of the two enzymes.

Correlation of Structure and Substrate Specificity. The previous studies on the IU-NH from *C. fasciculata* showed a broad substrate specificity with a preference for inosine and uridine, but 5-methyluridine, cytosine, adenosine, guanosine, and purine riboside were also substrates, with k_{cat}/K_M values in the range of 10^3-10^5 M^{-1} s^{-1}. In contrast, substitutions in the ribose led to complete or near complete loss of catalytic function. These results indicated a lack of

multiple specific interactions between the enzyme and the base but specific and multiple interactions with the ribosyl. This is confirmed by the proposed active site architecture. The bottom of the cavity contains multiple carboxylate residues; thus the nucleoside base is likely to be accommodated on the outer more hydrophobic part of the pocket. The walls of the cavity are largely constituted by hydrophobic residues (Phe, Ile, Met, Tyr). These residues provide a suitable environment to accommodate the nitrogenous base, via hydrophobic and $\pi-\pi$ stacking interactions. Few polar or charged groups are present that could be involved in specific interactions with the purine ring substituents in order to allow the enzyme to discriminate between the various nucleosides. In both OPRTase and HGPRTase the bound nucleotides are accommodated with the base deeply inserted in a niche of the molecular surface, stacked against the aromatic ring of a phenylalanine residue. In human HGPRTase, the purine ring is also in contact with the side chain of an isoleucine residue. Base specificity is provided in OPRTase by an arginine residue that forms specific hydrogen-bonding interactions with the O4 of orotate and the carboxylate on the pyrimidine ring interacting with a main chain nitrogen. In human HGPRTase, the bound GMP has a hydrogen-bonding pattern that explains the discrimination of the enzyme between GMP and AMP, IMP, and XMP. In uracil DNA glycosidase (Savva et al., 1995) the pyrimidine product of the reaction is found in a similar crevice, interacting with a phenylalanine and a tyrosine residue. The absolute discrimination of the enzyme for uracil over all other nitrogenous bases is ensured by a network of hydrogen bonds with residues that are highly conserved in the family of uracil DNA glycosidases.

The molecular surface of the IU-NH shows no niche or crevice where the purine ring could be accommodated. The lack of base specificity of IU-NH, completely different from the three above-mentioned examples, is readily explained by a model where the enzyme derives most of its catalytic power by interaction with the ribosyl moiety of the nucleoside. The high specificity of the enzyme for the ribose group, with all three hydroxyls required for efficient catalysis, suggests that the carboxylate cluster provides hydrogen-bonding recognition, with ribose bound in the bottom of the pocket and the base interacting with the side chains of the upper, hydrophobic region.

Structure of the Catalytic Site and Implications for the Catalytic Function. An applied interest in IU-nucleoside hydrolase is related to the requirement for purine salvage by parasitic protozoa and the novel use of nucleoside hydrolases in these organisms. Mammalian purine nucleoside phosphorylase and protozoan IU-NH differ in terms of substrate specificity, catalytic mechanism, and three-dimensional structure. However, calf spleen PNP is also capable of a slow hydrolysis of inosine to hypoxanthine and ribose (Kline & Schramm, 1992) but with a mechanism that differs from that for IU-NH. Purine nucleoside phosphorylase activates the leaving group by protonation of the purine base and neutralizes the developing positive charge of the oxycarbonium transition state with the negatively charged phosphate; the enzyme-bound attacking nucleophile (Kline & Schramm, 1995). The common catalytic feature of the *N*-ribohydrolases, glucosidases, and phosphorylases is the oxycarbonium character at the transition state [e.g. Sinnot (1990)]. This transition state can be achieved either by protonation of the leaving group (as described for PNP), in a classic general acid catalysis, or by the catalytic site interactions with the ribose that induce oxycarbonium character by distortion of the ring, as proposed for lysozyme (Gorenstein et al., 1977) and more recently for the enzyme "hydrogen-forming" methylenetetrahydromethanopterin dehydrogenase (Thauer, 1995). Recent data have demonstrated that *p*-nitrophenyl β-D-riboside is a better substrate for IU-NH than the natural substrate, inosine, despite the fact that the leaving *p*-nitrophenyl group is a poor candidate for activation by enzymatic protonation (Mazzella et al., 1996). These results indicate that IU-NH derives most of its catalytic efficiency by geometric and electronic distortion of the ribose ring toward the oxycarbonium transition state. Distortion of the ribose ring by hydrogen-bonding interactions to the carboxylates presumably allows the molecular orbitals to orient in a geometry which favors the hyperconjugation necessary for oxycarbonium ion generation. As positive charge builds on the oxycarbonium ion to form the transition state, it is stabilized by interaction with the nearby carboxylate(s). Consequently, the C1′ atom of the ribose rehybridizes toward sp^2, and the N-glycosidic bond is elongated. An enzyme-directed water molecule approaches the C1′ atom, forming the product α-ribose, and the nitrogenous base is protonated, allowing the release of the products.

Studies of the pH−rate profiles and transition state analysis have provided information about the groups provided by the enzyme for catalysis (Parkin & Schramm, 1995; Horenstein & Schramm, 1993a). An unprotonated group of pK_a 7.1 and a protonated residue, pK_a of 9.1, are required for catalysis. The unprotonated residue now seems likely to be one (or more) of the aspartic acids, while the protonated residue was proposed to be a lysine or histidine. From the structure of the proposed active site, four aspartate residues, Asp 10, Asp 14, Asp 15, and Asp 242, are at the bottom of the cavity. The first, third, and fourth are pointing toward the center of the cavity and are involved in chelating an electron-dense ion in both subunits of the asymmetric unit. Although there has been no demonstrated requirement of metal ions for IU-NH activity, competitive inhibition by cationic buffers has been observed. These observations indicate that one or more of the carboxylates are anionic, providing centers for ionic interactions. Asp 14 is involved in hydrogen-bonding interactions with His 241. One of these aspartic acid residues is proposed to be the ionizable group that will act as a general base, extracting the proton from the incoming water nucleophile. The proposed role of the other carboxyl groups is to bond and orient the ribosyl hydroxyls and to stabilize the positive charge in the transition state. The residue His 241 is a candidate to act as the protonating agent for the leaving nitrogenous base. A clear answer to the enzyme−substrate contacts in the transition state may be provided by the studies now underway to provide the three-dimensional structure of the complex between IU-NH and phenyliminoribitol derivatives. These compounds have been proposed to be transition state inhibitors (Horenstein & Schramm, 1993b) and demonstrate K_i values of 5−300 nM.

A comparison of unidentified open reading frames with similar deduced amino acid sequences to IU-NH is shown in Table 3. The residues which surround the catalytic site of IU-NH are marked and are conserved in all six of these sequences. The functions of the proteins in Table 3 were previously unknown. This comparison with IU-NH suggests

Table 3: Comparison of Active Site Residues with Sequences Similar to IU-Nucleoside Hydrolase from *C. fasciculata*[a]

```
          1       ↓ ↓ ↓↓↓ ↓↓↓  ↓↓                ↓↓↓  ↓       50
  Lmajor_b ...MPRKIIL DCDPGIDDAV AIFLAHGNPE IELLAITTVV GNQILEKVTQ
ybp2_desam ...MPRYAII DSDTASDDTI AILLASKF.. FKLLGITIVA GNVKFENEIK
  Cf_IU-NH ...MPKKIIL DCDPGLDDAV AILLAHGNPE IELLAITTVV GNQTLAKVTR
 schz_pomb .....MKIII DTDPGQDDAI TALLAIASPE IELLGVTTVA GNVPVSMTTR
   S_cerev MTVSKIPIWL DCDPGHDDAI AILLGCFHPA FNLLGISTCF GNAPPENTDY
 Ec_Lspdap ...MRLPIFL DTDPGIDDAV AIAAAIFAPE LDLQLMTTVA GNVSVEKTTR
   EcoHU47 ..MEKRKIIL DCDPGHDDAI AIMMAAKHPA IDLLGITIVA GNQTLDKTLI
          51                                                  100
  Lmajor_b NARLVA*... ......... ......... ......... .........
ybp2_desam NALFTVEYFN L.D.VPVFIG SSRPIMGKWS TVEEVHGNNG IGDWKIEEP.
  Cf_IU-NH NAQLVADIAG ITG.VPIAAG CDKPLVRKIM TAGHIHGESG MGTVAYPAEF
 schz_pomb NALQMLDLAG RPD.IPVYAG SNKPLLRAPI TATHVHGASG FEGAVLPPP.
   S_cerev NARSLLTAMG KAQAIPVYKG AQRPWKREPH YAPDIHGISG LDGTSLLPKP
 Ec_Lspdap NALQLLHFWN .AE.IPLAQG AAVPLVRAPR DAASVHGESG MAGYDFVEHN
   EcoHU47 NGLNVCQKLE I.N.VPVYAG MPQPIMRQQI VADNIHGETG LDGPVFEP.L
          101                    ↓   ↓ ↓                  150
  Lmajor_b ......... ......... ......... ......... .........
ybp2_desam KISPEKEHAI DAIIRLSKEY EGE.LEILAV SPLTNLALAY LKDPTIVKRI
  Cf_IU-NH KNKVDERHAV NLIIDLVMSH EPKTITLVPT GGLTNIAMAA RLEPRIVDRV
 schz_pomb SRKENEGHAV DFIIDTLRNN EPGTITICTI GPLTNIALAL NKAPEVIQRA
   S_cerev TFEARTDKTY IEAIEEAILA NNGEISFVST GALTTLATVF RCKPYLKKSV
 Ec_Lspdap RKPLGIPAFL AIRDALMRAP EP..VTLVAI GPLTNIALLL SQCPECKPYI
   EcoHU47 TRQAESTHAV KYIIDTLMAS DGD.ITLVPV GPLSNIAVAM RMQPAILPKI
          151 ↓↓↓                  ↓↓↓↓                  200
  Lmajor_b ......... ......... ......... ......... .........
ybp2_desam KKVWIMGGAF SR.GNTTP.. IAEFNFWVDP EAA..... .........
  Cf_IU-NH KEVVLMGGGY HE.GNATS.. VAEFNIIIDP EAAHIVFNES ...WQVTMVG
 schz_pomb KQIVMMAGAF SEVGNITP.. AAEFNIYVDP HAAQMVLSSG ...IPIVMMP
   S_cerev KYISIMGGGL HGLGNCNPNL SAEFNVWIDP DAANYIFRDP DVKDKCIVVP
 Ec_Lspdap RRLVIMGGSA GR.GNCTP.. NAEFNIAADP EAAACVFRSG ...IEIVMCG
   EcoHU47 REIVLMGGAY GT.GNFTP.. SAEFNIFADP EAARVVFTSG ...VPLVMMG
          201                                                 250
  Lmajor_b ......... ......... ......... ......... .........
ybp2_desam ↓........ ......... ......... ......... .........
  Cf_IU-NH LDLTHQALAT ...PPILQRV KEVDTNPARF MLEIMDYYTK IYQSNRYMAA
 schz_pomb LDITHQLHTS ...AKRIARM EALPNRVGPV VAAWLRMEKA YEAKKYGTDG
   S_cerev LNLTHKAIAT YKVNEMIYNE KNNSKLRELF LELFQFFAHT YKDMQGFESG
 Ec_Lspdap LDVTNQAILT ...PDYLSTL PQLNRTGKML HALFSHYRSG SMQS.....G
   EcoHU47 LDLTNQTVCT ...PDVIARM ERAGGPAGEL FSDIMNFTLK TQFENYGLAG
          251                                                 300
  Lmajor_b ...↓↓..... ......... ......... ......... .........
ybp2_desam ......... ......... ......... ......... .........
  Cf_IU-NH AAVHDPCAVA YVID..... ...PSVMTTE RVPVDIELTG KLTLGMTVAD
 schz_pomb GPLHDPNTVM WLLR..... ...PDIYSGR KVNVQIETQS ELTMGMSVVD
   S_cerev PPIHDPVALM PLLEFYGWDP SSAVGFRYKR MDISCIDDVF NENSGKIIIE
 Ec_Lspdap LRMHDLCAIA WLVR..... ...PDLFTLK PCFVAVETQG EFTSGTTVVD
   EcoHU47 GPVHDATCIG YLIN..... ...PDGIKTQ EMYVEVDVNS GPCYGRTVCD
          301                      338
  Lmajor_b ......... ......... ......... ........
ybp2_desam ......... ......... ......... ........
  Cf_IU-NH FRNPRPEHCH TQVAVKLDFE KFWGLVLDAL ERIGDPQ*
 schz_pomb WWQVGLLPAN VTFLRTVDDD EFYEVLIERL GRLP*...
   S_cerev KEYPNDSDVG TIIGLDLNIQ YFWDQIFEAL NRA.....
 Ec_Lspdap IDGCLGKPAN VQVALDLDVK GFQQWVAEVL ALAS....
   EcoHU47 ELGVLGKPAN TKVGITIDTD WFWGLVEECV RGYIKTH.
```

[a] Alignment with pileup (GCG package). Abbreviations: Lmajor _b, a partial cDNA sequence from *Leishmania major* corrected for one nucleotide base and translated; ybp2_desam, unassigned open reading frame (orf) from the *Desulfurolobus ambivalens* bps2 gene region; Cf_IU-NH, IU-nucleoside hydrolase from *C. fasciculata*; schz_pomb, unassigned orf from *Schizosaccharomyces pombe* chromosome II; S_cerev, unassigned orf from the *Saccharomyces cerevisiae* chromosome IV right arm; Ec_lspdap, unassigned orf from the Lsp-dapB intergenic region from *E. coli*; EcoHU47, unassigned orf from the *E. coli* nfo-FruA intergenic region. The numbering scheme is according to the open reading frame from S_cerev. Residues within 6 Å of the proposed active site and conserved among 66% or more of the similar sequences are indicated by ↓.

that they may have similar catalytic function and thus that the nucleoside hydrolases are widely distributed in bacteria and the lower eukaryotes. It is significant that no homologues have been identified in mammalian sequences.

Conclusions. The three-dimensional structure of inosine−uridine nucleoside hydrolase from *C. fasciculata* is the first of the nucleoside *N*-ribohydrolases to be characterized from the protozoa. The structure is novel in its protein topology and in the presence of a carboxylate cluster lining the floor of the presumptive catalytic site pocket. The presence of a bound metal ion neutralizes part of the negative charge of the carboxyls. The carboxyls are arranged in a cluster which could interact with all hydroxyls of the ribosyl of a nucleoside as well as provide an electronic environment to stabilize a developing oxycarbonium. The portion of the catalytic site proposed to interact with the purine leaving group is hydrophobic with a single ionizable histidine. Amino acids which form the putative catalytic site are represented in at least six other proteins on the basis of their presence in unidentified open reading frames from other microorganisms.

ACKNOWLEDGMENT

The authors thank Drs. S. C. Almo and S. L. Roderick for many helpful discussions.

REFERENCES

Adams, M. J., Ford, G. C., Koekoek, R., Lentz, P. J., McPherson, A., Rossmann, M. G., Smiley, I. E., Schevitz, R. W., & Wonacott, A. J. (1970) *Nature 227*, 1098−1103.

Bränden, C. I. (1980) *Q. Rev. Biophys. 13*, 317−338.

Brünger, A. T. (1992) *X-PLOR Version 3.0 Manual: a System for Crystallography and NMR*, Yale University, New Haven, CT.

Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 223*, 458−460.

Cross, G. A. M. (1975) *Parasitology 71*, 393−417.

Dewey, V. C., & Kidder, G. W. (1973) *Arch. Biochem. Biophys. 157*, 380−387.

Eads, J. C., Scapin, G., Xu, Y., Grubmeyer, C., & Sacchettini, J. C. (1994) *Cell 78*, 325−334.

Ealick, S. E., Rule, S. A., Carter, D. C., Greenhough, J. J., Babu, Y. S., Cook, W. J., Habsch, J., Helliwell, J. R., Stoeckler, J. D., Parks, R. E., Jr., Chen, S., & Bugg, C. E. (1990) *J. Biol. Chem 265*, 1812−1820.

Estupiñán, B., & Schramm, V. L. (1994) *J. Biol. Chem. 269*, 23068−23073.

Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.

Furey, W., & Swaminathan, S. (1990) *PHASES: A Program Package for the Processing and Analysis of Diffraction Data from Macromolecules, Am. Cryst. Assoc. Annu. Mtg. Program Abstr. 18*, 73.

Gopaul, D. N., Meyer, S. L., Degano, M., Sacchettini, J. C., & Schramm, V. L. (1996) *Biochemistry 35*, 5963−5970.

Gorenstein, D. G., Findlay, J. B., Luxon, B. A., & Kar, D. (1977) *J. Am. Chem. Soc. 99*, 3473−3478.

Hammond, D. J., & Gutteridge, W. E. (1984) *Mol. Biochem. Parasitol. 13*, 243−261.

Horenstein, B. A., & Schramm, V. L. (1993a) *Biochemistry 32*, 7089−7097.

Horenstein, B. A., & Schramm, V. L. (1993b) *Biochemistry 32*, 9917−9925.

Horenstein, B. A., Parkin, D. W., Estupiñán, B., & Schramm, V. L. (1991) *Biochemistry 30*, 10788−10795.

Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., & Salemme, F. R. (1987) *J. Appl. Crystallogr. 20*, 383−387.

Jones, T. A. (1985) *Methods Enzymol. 115*, 157−171.

Kidder, G. W., Dewey, V. C., & Nolan, L. L. (1978) *J. Cell Physiol. 96*, 165−170.

Kline, P. C., & Schramm, V. L. (1992) *Biochemistry 31*, 5964−5973.

Kline, P. C., & Schramm, V. L. (1995) *Biochemistry 34*, 1153−1162.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946.

Krenitsky, T. A. (1967) *Mol. Pharmacol. 3*, 526−536.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.

Mazzella, L. J., Parkin, D. W., Tyler, P. C., Furneaux, R. H., & Schramm, V. L. (1996) *J. Am. Chem. Soc. 118*, 2111−2112.

Miller, R. L., Sabourin, C. L. K., Krenitsky, T. A., Berens, R. L., & Marr, J. J. (1984) *J. Biol. Chem. 265*, 5073−5077.

Neubert, T. A., & Gottlieb, M. (1990) *J. Biol. Chem. 259*, 7236−7242.

Nicholls, A., Sharp, K., & Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281.

Parkin, D. W., & Schramm, V. L. (1995) *Biochemistry 34*, 13961−13966.

Parkin, D. W, Horenstein, B. A., Abdulah, D. R., Estupiñán, B., & Schramm, V. L. (1991) *J. Biol. Chem. 266*, 20658−20665.

Pizzi, T., & Taliaferro, W. H. (1960) *J. Infect. Dis. 107*, 100−107.

Read, R. J. (1986) *Acta Crystallogr., Sect. A 42*, 140−149.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167−339.
Rossmann, M. G., & Blow, D. M. (1962) *Acta Crystallogr. 15*, 24.
Savva, R., McAuley-Hecht, K., Brown, T., & Pearl, L. (1995) *Nature 373*, 487−493.
Scapin, G., Grubmeyer, C., & Sacchettini, J. C. (1994) *Biochemistry 33*, 1287−1294.
Schramm, V. L., Horenstein, B. A., & Kline, P. C. (1994) *J. Biol. Chem. 269*, 18259−18262.
Sinnot, M. L. (1990) *Chem. Rev. 90*, 1171−1202.
Smith, J. L., Zaluzec, J., Wery, J. J.-P., Niu, L., Switzer, R. L., Zalkin, H., & Satow, Y. (1994) *Science 264*, 1427−1433.
Steiger, R. F., & Steiger, E. (1977) *J. Protozool. 24*, 437−441.
Thauer, R. K. (1995) *Angew. Chem. 107*, 2418.
Tong, L., & Rossmann, M. G. (1990) *Acta Crystallogr., Sect. A 46*, 783−790.
Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1988) *Acta Crystallogr., Sect. A 43*, 489−501.
Tuttle, J. V., & Krenitsky, T. A. (1980) *J. Biol. Chem. 255*, 909−916.
Vellieux, F. M. D. (1993) *DEMON Program Suite Manual*, Grenoble, France.
Vickerman, K. (1985) *Br. Med. Bull. 41*, 105−114.
Wang, B. C. (1985) *Methods Enzymol. 115*, 90−112.

BI952999M