Inosine−Uridine Nucleoside Hydrolase from *Crithidia fasciculata*. Genetic Characterization, Crystallization, and Identification of Histidine 241 as a Catalytic Site Residue[†,‡]

Deshmukh N. Gopaul, Sheryl L. Meyer,[§] Massimo Degano,[||] James C. Sacchettini,* and Vern L. Schramm*

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

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

ABSTRACT: Protozoa depend on purine salvage for nucleic acid synthesis. An abundant salvage enzyme in *Crithidia fasciculata* is the inosine−uridine nucleoside hydrolase (IU-nucleoside hydrolase). The enzyme was cloned by polymerase chain reaction techniques using primers corresponding to the amino acid sequences of tryptic fragments and to the miniexon of *C. fasciculata*. The full-length cDNA was expressed in *Escherichia coli* and the protein purified to >99% homogeneity. The open reading frame encodes a protein of 315 amino acids. Enzyme purified from *C. fasciculata* was missing the N-terminal Met and gave a major mass peak of 34 194 amu by mass spectrometry. Predicted mass from the DNA sequence for the Met-processed enzyme was 34 196. A pET3d-IUNH construct expressed in *E. coli* introduced MetAla instead of MetPro at the N-terminus. Enzyme purified from this construct also had a processed N-terminus and gave predicted and observed masses of 34 168 and 34 170 amu, respectively. The amino acid sequence for IU-nucleoside hydrolase has no close relatives among the known proteins. A cDNA clone of unknown function from *Leishmania major* shows near identity in the N-terminal deduced amino acid sequence. Open reading frames near 1 and 47 min on the *E. coli* chromosome and from two yeast genomes encode for proteins of similar size with substantial amino acid identity. Mutation of His241Ala caused a 2100-fold loss in k_{cat} for inosine but a 2.8-fold increase in k_{cat} with *p*-nitrophenyl β-D-ribofuranoside, establishing the location of the catalytic site and implicating His241 as a proton donor for leaving group activation. IU-nucleoside hydrolase from *C. fasciculata* and the protein expressed in *E. coli* were crystallized and diffract to 2.5 and 2.1 Å resolution, respectively. Both belong to the $P2_12_12$ orthorhombic space group with unit cell parameters $a = 63.5$ Å, $b = 131.9$ Å, $c = 90.1$ Å, and $\alpha = \beta = \gamma = 90°$. Two subunits of the tetrameric enzyme are present in the asymmetric unit. The following paper reports the X-ray crystal structure for this enzyme.

Trypanosomes are protozoan parasites that are the causative agents of several epidemiologically significant afflictions including Chagas disease and sleeping sickness. The protozoan parasites are deficient in *de novo* synthesis of purines and therefore rely on salvage pathways. Nucleoside hydrolases have been implicated in the salvage pathway of purine nucleosides in the trypanosome, *Crithidia fasciculata* (Parkin et al., 1991; Estupiñán & Schramm, 1994). This organism provides a convenient source of trypanosomal enzymes, since it does not infect mammals but has a GC-rich genome which is similar in base composition to *Trypanosoma* and *Leishmania* species (Alonso et al., 1992). The most abundant nucleoside hydrolase in *C. fasciculata* prefers inosine and uridine as substrates (IU-nucleoside hydrolase)[1] but catalyzes the hydrolysis of all of the commonly occurring purine and pyrimidine nucleosides into ribose and the associated base. The enzyme is a homotetramer of 34 kDa subunits and exhibits no known allosteric properties (Parkin et al., 1991). Extracts from a variety of protozoan parasites indicate the presence of several nucleoside hydrolases with differing substrate specificities. In *C. fasciculata*, a second enzyme is present which is highly active with guanosine and inosine but is nearly inert with pyrimidine nucleosides and adenosine (GI-nucleoside hydrolase; Estupiñán & Schramm, 1994). Three nucleoside hydrolases have been reported from *Leishmania donovani* (Koszalka & Krenitsky, 1979), and four have been reported from *Trypanosoma cruzi* (Miller et al., 1984). None of these catalytic activities have been identified in mammals, suggesting that the nucleoside hydrolases play specific roles in the protozoa and may be reasonable targets for antibiotic design (Horenstein & Schramm, 1993b). Structural and metabolic studies on these enzymes have suffered from the lack of genetic information. As a prelude to structural determination, the gene for IU-nucleoside hydrolase has been cloned from *C. fasciculata*, sequenced, overexpressed in *Escherichia coli*, and crystallized in a form suitable for X-ray diffraction.

[†] This work was supported by Research Grant GM41916 from the National Institutes of Health. Data in this paper are from a thesis to be submitted in partial fulfillment of the requirements for the Ph.D. degree at the Albert Einstein College of Medicine, Yeshiva University.

[‡] The nucleotide sequence reported in Figure 1 has been submitted to the GenBank/EMBL Data Bank with accession number Bank It 26658 U43371.

* Corresponding authors. Fax: (718) 430-8565. Email: vern@aecom.yu.edu; sacchett@aecom.yu.edu.

[§] Present address: Cephalon Inc., 145 Brandwine Parkway, West Chester, PA 19380-4245.

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

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

[1] Abbreviations: IU-nucleoside hydrolase, nonspecific inosine−uridine nucleoside hydrolase from *Crithidia fasciculata*; nitrophenyl riboside, *p*-nitrophenyl β-D-ribofuranoside; HPLC, high-performance liquid chromatography; PCR, polymerase chain reaction; bp, base pairs; PMSF, phenylmethanesulfonyl fluoride; pET3d-IUNH, the pET3d plasmid (Novagen Inc.) containing the IU-nucleoside hydrolase gene inserted in the *Nco*I−*Bam*HI restriction sites.

S0006-2960(95)02998-9 CCC: $12.00

The chemical mechanism and transition state structure for IU-nucleoside hydrolase have been determined by kinetic isotope effect measurements and shown to proceed through an oxycarbonium-like transition state (Horenstein et al., 1991). Analogues which resemble the transition state in charge and geometry are powerful inhibitors of the enzyme (Horenstein & Schramm, 1993a,b; Boutellier et al., 1994; Horenstein et al., 1993). The amino acids which stabilize the transition state and form favorable contacts with tight-binding inhibitors of the enzyme can best be examined by determination of its three-dimensional structure. This could provide useful information to explain the catalytic mechanism and the tight binding of transition state inhibitors and to assist in the design of antitrypanosomal antibiotics. This work reports the cloning of IU-nucleoside hydrolase, determination of its primary amino acid sequence, its expression in *E. coli*, a purification protocol for the recombinant enzyme, crystallization of the enzyme, and a site-directed mutant which identifies the catalytic site.

METHODS

Initial Rate and Alternative Substrate Studies. Hydrolysis of inosine to hypoxanthine and ribose was monitored by the release of reducing sugar or by continuous recording of optical absorbance at 280 nm (Parkin et al., 1991). Hydrolysis of *p*-nitrophenyl β-D-ribofuranoside (nitrophenyl riboside) was followed by release of the *p*-nitrophenylate ion at 400 nm, in the indicated buffer (Mazzella et al., 1996). The extinction coefficient for the *p*-nitrophenylate ion under specific conditions was determined by the absorbance at 400 nm following complete conversion to products.

Peptide Isolation and Sequencing. IU-nucleoside hydrolase was purified from *C. fasciculata* (Parkin et al., 1991) and digested with L-1-(tosylamido)-2-phenylethyl chloromethyl ketone-treated trypsin or cyanogen bromide. Peptide fragments were isolated by reverse-phase HPLC, eluting with a gradient of acetonitrile in 0.1% trifluoroacetic acid. Amino acid sequences were obtained by automated Edman gas-phase sequencing from the intact protein (N-terminal sequence) and from several internal peptides. The amino acid sequence obtained from the N-terminus was PKKIILDxDPGLDDAVAILLAHGNPEIELLAITTGVGNQTLAKVTLNVQL. Tryptic and cyanogen bromide-derived peptides gave the following amino acid sequences which were of variable quality: FMTAGYIAGSGMGTVAYPAGFKK, GTVAYPAEFKNKVDERHAVN, VREPKTITLVPRYGLTNIA, VGLDLTHQALATPNILQRVK, AARLEPRIVDRVKEV, TAGHIHGESGT, and GYYTKIYQSNRYT.

DNA Primers. The N-terminus and an internal amino acid sequence were reverse transcribed into DNA sequences, and PCR primers were designed using the codon bias reported for *C. fasciculata* (Alonso et al., 1992). PCR reactions using the sequences 5′-GACCCGGGCCTGGACGACGCGGTGGCG-3′ for a primer in the N-terminal region (DPGLDDAVA) and 5′-CTGGTA(G,A)ATCTT(C,G)GTGTAGTA-3′ as an internal primer (YYTKIYQ) with genomic DNA from *C. fasciculata* as template gave a PCR product of approximately 650 base pairs. DNA sequencing indicated that this PCR product contained the coding region for the expected N-terminal region of the protein as well as several internal peptides which had been identified by direct sequencing. This segment was used to design oligonucleotide primers to isolate the gene from cDNA by PCR.

Isolation of the cDNA and Genomic DNA for IU-Nucleoside Hydrolase. Poly(A) mRNA was isolated from extracts of *C. fasciculata* on an oligo(dT) column (Invitrogen). The mRNA was reverse transcribed into cDNA by reverse transcriptase (Boeringher Mannheim, Maloney murine leukemia virus) and using randomly generated hexameric oligonucleotides (Perkin-Elmer). This cDNA was used as a template for further PCR selection using *Taq* polymerase (Perkin-Elmer). The 5′-region was cloned by PCR amplification using the synthetic sense primer (5′-GCTATATAAGTATCAGTTTCTGTAC-3′) for the miniexon-derived sequence common to all *C. fasciculata* mRNA (Muhich et al., 1987) and the codon-biased antisense primer (5′-CGTCCACCTTGTTCTTGAACTC-3′) for the internal peptide EFKNKVDE. The 3′-region was cloned using the sense primer (5′-CTTCATGTTGGAGATCATGGAC-3′) to an internal peptide sequence (FMLEIMDY) and a 16-mer oligo(dT) against the poly(A) tail. The PCR products were ligated into the TA vector (Invitrogen) and sequenced (Sequenase 2.0, USB). Fragments of 450 and 900 bp were obtained for the 5′- and 3′-regions, respectively. Primers 5′-AAAACCATGGCGAAGAAGATCATCCTC-3′ and 5′-GAATTCGGGATCCTGTTACTG-3′ were based on the nucleotide sequence of the 5′- and 3′-regions and were designed to incorporate *Nco*I and *Bam*HI sites, respectively. These were used to amplify the full-length gene by PCR using genomic DNA from *C. fasciculata* as template. The dsDNA was ligated into the pET-3d vector (Novagen Inc.) to give the pET3d-IUNH construct. Both DNA strands were sequenced using overlapping primers.

Site-Directed Mutagenesis of His241Ala in IU-Nucleoside Hydrolase. Primers spanning position 241 were designed in both sense and antisense directions with modifications for the codon of His241 (CAC) to Ala (GCC). Two additional primers were synthesized corresponding to the unique restriction sites *Sac*II and *Bam*HI located 300 bp upstream and 230 bp downstream, respectively, from the mutation site. A PCR reaction with the Pwo DNA polymerase (Boehringer Mannheim) generated an upstream region from the pET3d-IUNH dsDNA plasmid using the sense *Sac*II site primer with the antisense His241Ala mutation, and a downstream region using the antisense primer containing a *Bam*HI site and the sense His241Ala primer. The PCR products were purified from a 1% agarose gel using the Qiaquick gel extraction kit (QIAGEN Inc., Chatsworth, CA) and used as template in a second series of PCR with the *Sac*II and *Bam*HI site primers. The resultant PCR product was ligated into the pET3d-IUNH construct and sequenced to verify the mutation. The mutated plasmid was transformed into BL21(DE3)pLys S *E. coli* cells (Novagen Inc., Madison, WI) for protein expression. Protein was purified using the protocol described for the recombinant enzyme. The mutant protein was detected by size on SDS−PAGE gels. The purified protein was further characterized by electrospray ionization mass spectrometry.

Purification of Recombinant IU-Nucleoside Hydrolase. The 0.95 kb open reading frame for IU-nucleoside hydrolase was cloned into the pET3d (Novagen) expression vector using *Nco*I and *Bam*HI restriction sites. This construct was transformed into BL21(DE3)pLys S (Novagen) *E. coli* cells. Cells were grown in Luria−Bertani broth supplemented with 50 μg/mL carbenicillin. Cell cultures were grown in shaken flasks for 6−8 h at 30 °C. The cells were harvested by centrifugation and resuspended in 20 mM Tris-HCl, pH 7.4,

1 mM dithiothreitol, 1 mM ethylenediaminetetraacetic acid, 15 μM PMSF, 10 μg/mL *o*-phenanthroline, 0.5% Nonidet P-40, and 50 mM NaCl. The cells were stored and frozen at −70 °C.

Frozen cells were thawed, warmed to room temperature, and submitted to the French press or to three to five cycles of sonication at 80% output (Branson, 600 W sonicator) for 30 s with intermediate cooling cycles of 1 min. The cell lysate was centrifuged at 15 000 rpm for 40 min. Protamine sulfate (10% solution in H_2O) was added slowly, with constant stirring to a final concentration of 1% w/v. After centrifugation, supernate was fractionated with $(NH_4)_2SO_4$, and the pellet precipitating between 45, and 55% saturation was dissolved in a minimum volume of 20 mM triethanolamine, pH 7.8, containing 15 μM PMSF (the buffer) and dialyzed overnight at 4 °C against the buffer. The dialysate was applied to a FastQ anion-exchange chromatography column (150 mL bed volume) equilibrated with the buffer and eluted with a 600 mL linear buffered gradient of 0−0.5 M NaCl. The fractions containing nucleoside hydrolase protein and activity were pooled and dialyzed overnight at 4 °C against the buffer adjusted to pH 6.5. The dialysate was applied to a RedA column (Amicon, 75 mL bed volume) equilibrated in the buffer at pH 6.5 and eluted with a pH gradient of the buffer from 6.5 to 8.5. The activity peak was pooled, concentrated by ultrafiltration, rapidly frozen, and stored at −70 °C.

Crystallization Studies. Purified IU-nucleoside hydrolase proteins were crystallized by the hanging-drop vapor diffusion method. Protein at 25−40 mg/ml in 10 mM potassium phosphate, 10 mM glycine, pH 7.3, or 20 mM triethanolamine, pH 7.3, was diluted with equal volumes of poly(ethylene glycol) 4000 (6−14% w/v) in 100 mM potassium citrate, pH 4.6, and placed on a silanized coverslip, inverted, and sealed over a reservoir containing 0.7 mL of the undiluted poly(ethylene glycol) solution. Final drop sizes of 7 and 10 μL were used. Crystals of both recombinant and cell-derived enzymes were relatively large, reaching a maximum size of 0.2 × 1 × 1.8 mm in diamond or flat trapezoid shapes. Crystals grown from recombinant IU-nucleoside hydrolase were typically thicker, commonly 0.3−0.4 mm in the smallest dimension.

Initial X-ray Characterization. Native IU-nucleoside hydrolase crystals typically diffracted to better than 2.5 Å while those obtained from recombinant enzyme diffracted to greater than 2.1 Å resolution. X-ray diffraction data sets were collected from single crystals of both native and recombinant enzyme on a Siemens X-1000 area detector coupled with a Rigaku RU-200 rotating anode X-ray generator. The wavelength used was the copper Kα radiation (1.5418 Å) selected with a graphite monochromator. The diffraction data were indexed and integrated using the program XENGEN v2.1 (Howard et al., 1987). The space group symmetry and unit cell dimensions were determined using the program PRECESS, which is part of the PHASES package (Furey & Swaminathan, 1990).

RESULTS AND DISCUSSION

DNA Sequence of IU-Nucleoside Hydrolase. The open reading frame for the cDNA of IU-nucleoside hydrolase spans 945 base pairs followed by a TAA stop codon with an untranslated region of approximately 200 bp between the

```
-60 ttctcctcttctatttgcaggcaatccactgtcgctctatatcttaaaaaaaaaaaaag
    ---------+---------+---------+---------+---------+---------+

    M  P  K  K  I  I  L  D  C  D  P  G  L  D  D  A  V  A  I  L
  1 atgccgaagaagatcatcctcgactgcgacccgggtctcgatgacgccgtggcgatcctg 60
    ---------+---------+---------+---------+---------+---------+

    L  A  H  G  N  P  E  I  E  L  L  A  I  T  T  V  V  G  N  Q
 61 ctggcgcacggcaaccccgagatcgagctgctggccatcacgacggtggtaggcaaccag 120
    ---------+---------+---------+---------+---------+---------+

    T  L  A  K  V  T  R  N  A  Q  L  V  A  D  I  A  G  I  T  G
121 acgctggcgaaggtcacccgcaacgcgcagcttgtggccgacatcgccggcattacgggt 180
    ---------+---------+---------+---------+---------+---------+

    V  P  I  A  A  G  C  D  K  P  L  V  R  K  I  M  T  A  G  H
181 gtccctattgctgccggctgcgacaagccgctggtgcgcaagatcatgacggcgggccac 240
    ---------+---------+---------+---------+---------+---------+

    I  H  G  E  S  G  M  G  T  V  A  Y  P  A  E  F  K  N  K  V
241 attcacggcgagtcgggcatgggcacagttgcctacccggccgagttcaagaacaaggtg 300
    ---------+---------+---------+---------+---------+---------+

    D  E  R  H  A  V  N  L  I  I  D  L  V  M  S  H  E  P  K  T
301 gacgagcgccacgcggtgaacctgatcatcgaccttgtgatgagccacgagcccaagacg 360
    ---------+---------+---------+---------+---------+---------+

    I  T  L  V  P  T  G  G  L  T  N  I  A  M  A  A  R  L  E  P
361 atcacgctggtgccgacgggtggactgacgaacattgcgatggctgcgcgtctggagccg 420
    ---------+---------+---------+---------+---------+---------+

    R  I  V  D  R  V  K  E  V  V  L  M  G  G  G  Y  H  E  G  N
421 cggatcgtggaccgcgtgaaggaggtcgtgctgatgggcggcggctaccacgagggcaat 480
    ---------+---------+---------+---------+---------+---------+

    A  T  S  V  A  E  F  N  I  I  I  D  P  E  A  A  H  I  V  F
481 gcgacgtctgttgccgaatttaacatcattatcgatcccgaagcagcgcacatcgtcttt 540
    ---------+---------+---------+---------+---------+---------+

    N  E  S  W  Q  V  T  M  V  G  L  D  L  T  H  Q  A  L  A  T
541 aacgagagctggcaggtgacgatggtgggcctcgacctcacgcaccaggcgctggccaca 600
    ---------+---------+---------+---------+---------+---------+

    P  P  I  L  Q  R  V  K  E  V  D  T  N  P  A  R  F  M  L  E
601 ccgccgattctgcagcgtgtgaaggaggtggacacgaacccggcgcgcttcatgttggag 660
    ---------+---------+---------+---------+---------+---------+

    I  M  D  Y  Y  T  K  I  Y  Q  S  N  R  Y  M  A  A  A  A  V
661 atcatggactactacacgaagatttaccaaagtaaccgctatatggcggctgcggcggtg 720
    ---------+---------+---------+---------+---------+---------+

    H  D  P  C  A  V  A  Y  V  I  D  P  S  V  M  T  T  E  R  V
721 cacgacccttgcgcggtggcttacgtgattgacccgtcggtgatgacgacggagcgagtg 780
    ---------+---------+---------+---------+---------+---------+

    P  V  D  I  E  L  T  G  K  L  T  L  G  M  T  V  A  D  F  R
781 ccggtggacatcgagctgaccggtaagctgacgctggggatgacggtggctgatttccgt 840
    ---------+---------+---------+---------+---------+---------+

    N  P  R  P  E  H  C  H  T  Q  V  A  V  K  L  D  F  E  K  F
841 aacccacgcccggagcactgccacacgcaggtggcggtgaagctggacttcgagaagttc 900
    ---------+---------+---------+---------+---------+---------+

    W  G  L  V  L  D  A  L  E  R  I  G  D  P  Q  *
901 tggggcttggtgttggacgctctggagcgtatcggcgaccccagtaacaggatctcgaa 960
    ---------+---------+---------+---------+---------+---------+

1021 gattcaatcggaaagaagagatggcgtcagtaaagcagac  1060
     ---------+---------+---------+---------+
```

FIGURE 1: Sequence of the cDNA encoding for the IU-nucleoside hydrolase from *C. fasciculata*. The region −60 to 0 is 5′ to the initiation Met codon which is numbered 1. The TAA stop codon is indicated by an asterisk, followed by the sequence which leads to the poly(A) tail of the mRNA.

miniexon site and the initiation codon (Figure 1). Amplification of genomic DNA using primers from the cDNA gave the same length DNA, confirming the lack of introns in this gene. The stop codon is followed by approximately 500 bp of untranslated nucleotides prior to the polyadenylate tail of the mRNA. The coding region is 62.2% GC, compared to the overall GC content of 64.9% in a sample of the eight *C. fasciculata* genes reported recently (Alonso et al., 1992).

Sequence comparison against the GenEMBL data base revealed significant similarity in the deduced amino acid sequences of several open reading frames with previously unknown functions (Figure 2). The strongest is from a *Leishmania major* cDNA clone which shows 82.7% identity in the DNA sequence for the 120 base pairs at the 5′-coding region of the cDNA. Translation of the *Leishmania* sequence compared to the N-terminal amino acid sequence for IU-nucleoside hydrolase from *C. fasciculata* indicates that 36 of the 40 amino acids are identical. Thus, this open reading frame establishes a homologue of the *Crithidia* enzyme in

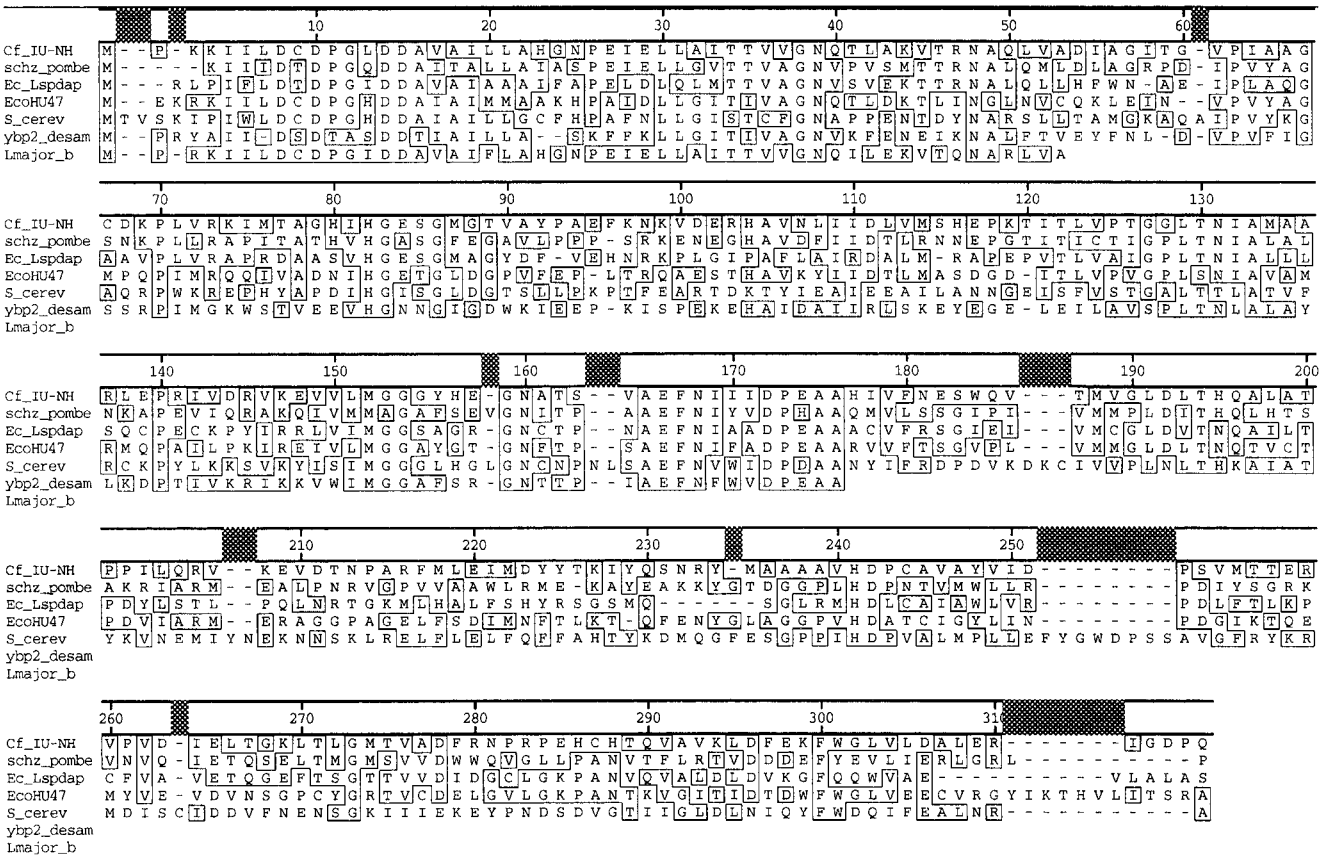

FIGURE 2: Alignment of amino acid sequences from IU-nucleoside hydrolase to translated amino acid sequences from unidentified open reading frames of DNA or cDNA. The Clustal method of DNASTAR was used with a stringency of 2.5 mutations/residue (PAM250) to align similar sequences. Comparisons are given for IU-nucleoside hydrolase (Cf_IU-NH) with the open reading frames identified from *S. pombe* (schz_pombe), from *E. coli* near the *lsp-dap*B intergenic region (Ec_Lspdap) and the *nfo-fru*A intergenic region (EcoHU47), from *S. cerevisiae* (S_cerev), from *Desulfurolobus ambivalens* sequence, bps2 gene, accession X64202 (ybp2_desam), and from a partial cDNA sequence from *L. major* (Lmajor_b). The boxed regions indicate amino acids with identity to the consensus sequence for this alignment. The shaded areas above the sequences are gaps introduced into the amino acid sequence of IU-nucleoside hydrolase from *C. fasciculata* to maintain maximum sequence alignment.

Leishmania. Significant similarities are also found in two open reading frames of unknown function in *E. coli*. The 40 N-terminal amino acid sequence of IU-nucleoside hydrolase was also used as a search strategy since many of the ionizable amino acids in the proposed catalytic site are composed of these sequences (Degano et al., 1996). An open reading frame located at 47 min on the *E. coli* chromosome in the *nfo-fruA* intergenic region encodes a protein of 33.7 kDa, which has identity in 28 of the first 40 amino acids, including all of the ionizable groups proposed at the catalytic site. A second sequence, located in the *lsp-dap*B intergenic region near min 1 of the *E. coli* chromosome, encodes for a protein of 32.6 kDa and shows 23 of the first 40 as identical amino acids near the N-terminus (Figure 2).

Two yeast genome open reading frames also show similarity to the IU-nucleoside hydrolase (Figure 2). Data from the GenEMBL data repository indicate a conserved sequence of 310 amino acids with 37.1% identity from *Schizosaccharomyces pombe*, identified by automatic sequencing of chromosome II in a cosmid library, accession number U33010. A second sequence with significant homology has been identified in *Saccharomyces cerevisiae*, also from systematic sequencing of the yeast genome. This sequence spans 257 amino acids and is 32.3% identical to IU-nucleoside hydrolase. It resides on the right arm of chromosome IV, accession number U32274. Significantly, all of these open reading frames place the presumptive catalytic site amino acid residues near the same regions of the deduced proteins and contain a His equivalent to His241 of IU-nucleoside hydrolase (Degano et al., 1996). This residue is proposed to be the proton donor for the leaving group in the hydrolysis of inosine (see below).

These results and previous reports of enzyme activity assays establish that the purine nucleoside N-ribohydrolase enzymes are well represented among the protozoa and yeast and in the bacteria. A chemically related reaction is catalyzed by the prokaryotic AMP nucleosidase which hydrolyzes the N-ribosidic bond of AMP (Leung et al., 1989). This nucleotide *N*-ribohydrolase resides at 43.3 min of the *E. coli* chromosome (Leung & Schramm, 1984), and catalyzes the reaction:

$$\text{AMP} + \text{H}_2\text{O} \rightarrow \text{adenine} + \text{ribose 5-phosphate}$$

There is no significant amino acid homology between IU-nucleoside hydrolase from *C. fasciculata* and the AMP nucleosidase from *E. coli*, despite the similarity in the chemical reactions. It is significant that no homologues of either enzyme are found in mammalian protein sequences, consistent with the proposal that nucleoside and nucleotide *N*-ribohydrolases are absent in mammalian pathways of purine metabolism and salvage.

Expression and Purification of Recombinant Nucleoside Hydrolase. E. coli cells transformed with the pET3d-IUNH

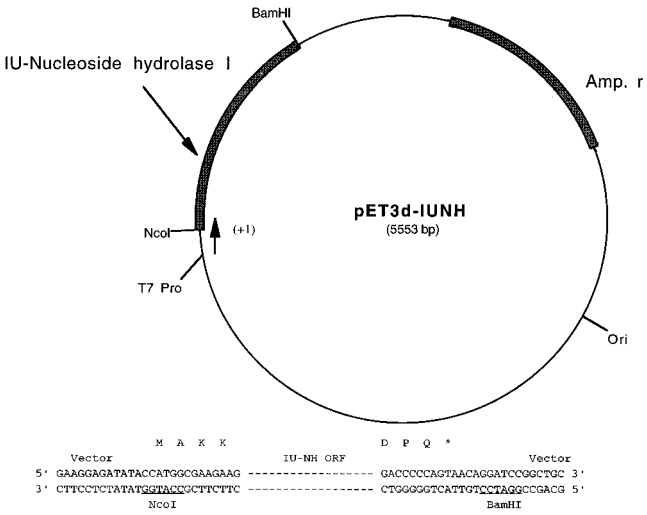

FIGURE 3: Plasmid for expression of IU-nucleoside hydrolase from *C. fasciculata*. The coding region is contained within the *Bam*HI−*Nco*I restriction fragment, placed in the homologous sites in pET3d. The base sequences surrounding the *Nco*I and *Bam*HI sites are shown. The open reading frame from the cDNA of native IU-nucleoside hydrolase is provided in Figure 1. In the expression vector, an Ala was introduced at the second codon to permit construction of the *Nco*I site. A *Bam*HI site was introduced near the native stop codon (TAA).

Table 1: Purification of Recombinant IU-Nucleoside Hydrolase from *C. fasciculata*, Expressed in *E. coli*

step	activity (μmol/min)	protein (mg)	specific activity (mol min^{-1} mg^{-1})	yield
protamine sulfate[a]	6970	1050	6.7	1.0
ammonium sulfate	4710	630	7.5	0.68
FastQ	2470	64	39	0.35
RedA	2440	32	76	0.35

[a] Ten grams of frozen cells was lysed in a French pressure cell and treated with a solution of 10% protamine sulfate to give a final concentration of 1%.

construct (Figure 3) and grown at 30 °C expressed 5−10% of the total soluble protein as IU-nucleoside hydrolase. Purification was simplified from that required for the enzyme from *C. fasciculata* where it represents approximately 0.1% of total soluble protein (Parkin et al., 1991). The improved protocol, summarized in Table 1, gave high purity enzyme in 35% yield (Figure 4) . N-Terminal amino acid sequencing of the purified protein gave the sequence AKKIILDCD-PGLDD which corresponded to the change in codon 2 from Pro to Ala for the recombinant IU-nucleoside hydrolase and confirmed the DNA sequence beyond the primer region. Gel filtration chromatography indicated that the protein eluted as a tetramer with M_r = 142 000, consistent with the tetrameric structure proposed earlier for this enzyme (Parkin et al., 1991). Further characterization was carried out by mass spectrometry, which gave a subunit $M_r = 34\,170 \pm 4$ compared to the expected value of 34 168 for the N-terminal Met-processed recombinant enzyme. A M_r of 34 194 ± 4 was determined for IU-nucleoside hydrolase isolated from *C. fasciculata* which corresponded to the expected M_r of 34 196 for the N-terminal Met-processed enzyme as deduced from the cDNA sequence (Figure 5). The close correspondence of observed and expected molecular weights for native and recombinant IU-nucleoside hydrolase and alignment of all peptide sequences indicate that the coding

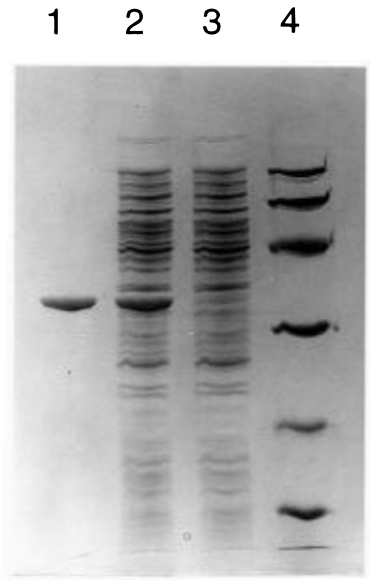

FIGURE 4: Expression of IU-nucleoside hydrolase in *E. coli*. Lane 4 contains the molecular weight markers: phosphorylase *b*, 106 000; bovine serum albumin, 80 000; ovalbumin, 47 500; carbonic anhydrase, 32 500; soybean trypsin inhibitor 27 500; and lysozyme, 18 500 M_r, respectively. Lane 3 contains extracts of the parental *E. coli* strain BL21(DE3) pLys S cells carrying the pET3d vector without an insert. Lane 2 is an extract from the same strain transformed with the plasmid containing the coding region for IU-nucleoside hydrolase. Lane 1 contains recombinant, purified IU-nucleoside hydrolase. Samples were denatured in 1% sodium dodecyl sulfate containing mercaptoethanol and run on SDS-denaturing polyacrylamide gels. Proteins are stained with Commassie brilliant blue.

region (Figure 1) provides reliable information for the protein.

Crystallization of Native and Recombinant IU-Nucleoside Hydrolases. IU-nucleoside hydrolase was crystallized using the hanging-drop vapor diffusion method in a variety of precipitant solutions. However, only the crystals grown from potassium citrate and poly(ethylene glycol) of 4000 average molecular weight were of good quality for high-resolution structural studies. The crystals of IU-nucleoside hydrolase purified from the trypanosome generally grew in thin plates, while those obtained from the recombinant enzyme were thicker and diffracted to slightly higher resolution. Crystals from native and recombinant protein are isomorphous and belong to the orthorombic crystal system. The unit cell dimensions are $a = 63.5$ Å, $b = 131.9$ Å, $c = 90.1$ Å, $\alpha = 90°$, $\beta = 90°$, and $\gamma = 90°$. A nonreduced X-ray diffraction data set was used to generate pseudoprecession photographs which were evaluated to establish the space group. Only the reflections with indices $h = 2n$ were observed along the a^* direction; similarly, only the reflections with indices $k = 2n$ were observed along the b^* direction. These observations suggested the presence of a 2-fold screw axis symmetry along both the crystallographic *a* and *b* axes. No systematic absences were detected along the c^* direction. On the basis of these observations, the space group was determined as being $P2_12_12$. For this preliminary characterization of the crystals, the data set was 92% complete to 2.7 Å resolution and had a *R*-factor of 7.8% for symmetrically equivalent reflections.

The unit cell volume is 754 353 Å^3, and since there are four equivalent positions in this space group, the asymmetric unit volume is 188 634 Å^3. Assuming two monomeric

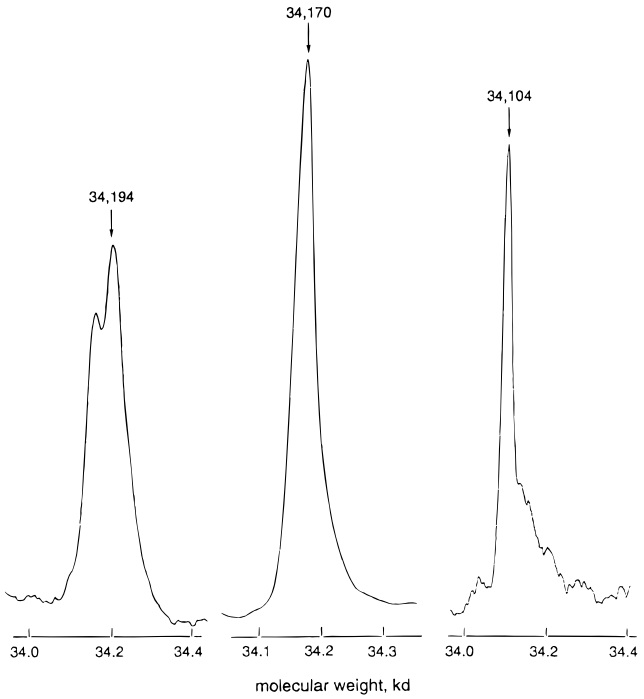

FIGURE 5: Electrospray mass spectrometry of native, recombinant, and His241Ala IU-nucleoside hydrolases. Samples of approximately 1 mg/mL of the purified enzymes were mixed with equal volumes of 50% acetonitrile−H_2O in 0.1% trifluoroacetic acid, acidified with 88% formic acid, and introduced at 2 μL/min into an PE-Sciex API-3 electrospray triple-quadrapole mass spectrometer typically run at 4000 V. The mass calibration was adjusted prior to introduction of the samples using poly(propylene glycol) m/z = 1000−2400. Spectra from left to right were obtained with IU-nucleoside hydrolases purified from *C. fasciculata* (major peak at 34 194 amu), recombinant Pro2Ala enzyme purified from *E. coli* (major peak at 34 170 amu), and Pro2Ala, His241Ala purified from *E. coli* (major peak at 34 104). The lower mass peak in the enzyme purified from *C. fasciculata* is 34 150 amu or 44 amu less than the major peak. The mass peak ratio varied among different enzyme preparations and is of unknown origin.

subunits in the asymmetric unit, and a protein density of 1.23 g/cm^3, a V_M of 2.75 Å^3/Da (Matthews, 1968) and a 56% solvent content in the unit cell are obtained. These values are in the typical range for protein crystals. These results are consistent with a situation where one of the tetrameric enzyme's molecular 2-fold axes is coincident with the space group's pure 2-fold axis, the crystallographic *c* axis.

Expression of His241Ala IU-Nucleoside Hydrolase. Site-directed mutagenesis of His241 was intended to establish the location of the catalytic site which was proposed from the nature of the protein surface in X-ray crystallographic studies (Degano et al., 1996) and the conservation of this residue in the presumed homologues of IU-nucleoside hydrolase (Figure 2). The protein was expressed in the pET3d construct and purified to near homogeneity using the method described in Table 1, but by following the protein band corresponding to IU-nucleoside hydrolase on SDS gel electrophoresis, since the His241Ala enzyme was inactive under standard assay conditions.

Kinetic Characterization of Recombinant IU-Nucleoside Hydrolase and His241Ala Enzyme. Placement of a convenient restriction site near the initiation codon required a change of the second codon from a proline to an alanine. Kinetic studies were used to establish the catalytic properties of the expressed and mutated enzymes compared to that isolated from *C. fasciculata* (Table 2). The k_{cat} for both inosine and nitrophenyl riboside is slightly greater with the recombinant enzyme than for the native. This may reflect the simplified purification procedure used for the recombinant enzyme, leading to a higher specific activity. The K_m values with inosine as substrate are similar for the native and recombinant enzymes. When nitrophenyl riboside (Mazzella et al., 1996) was used as substrate, the K_m obtained for recombinant enzyme is 2-fold less than for the native enzyme. These data indicate that Pro2Ala causes relatively minor changes in the catalytic properties of recombinant IU-nucleoside hydrolase.

Under standard assay conditions with inosine as substrate, the enzyme with His241Ala was inactive. At elevated enzyme concentration, the His241Ala mutant catalyzed inosine hydrolysis with a k_{cat} of 0.021 s^{-1}, a decrease of 2100-fold compared to the recombinant enzyme (Table 2). The K_m value was elevated by a factor of 3.5-fold, indicating that the affinity for substrate is only modestly altered and that the effect of the mutation is primarily on the catalytic turnover number. This result is consistent with the pH profile of the native enzyme which established that substrate binding is nearly independent of the enzymatic ionizable groups (K_m independent of pH) but that k_{cat} and the binding of a transition state inhibitor required a proton donor with a pK_a near 9.1 and a proton acceptor with a pK_a near 7.1 (Parkin & Schramm, 1995). On the basis of its location in the proposed catalytic site (Degano et al., 1996), His241 is a candidate for the group of pK_a 9.1 which protonates the hypoxanthine leaving group, a step which occurs to cause transition state formation (Horenstein et al., 1991). The transition state structure derived from kinetic isotope effects proposed N7 protonation of hypoxanthine (Figure 7), but protonation at any heteroatom of the hypoxanthine would convert the purine base to an improved leaving group. If His241 accomplishes this role, a substrate with a better leaving group would be expected to retain activity for the His241Ala enzyme.

Assay of His241Ala IU-nucleoside hydrolase with nitrophenyl riboside gave a k_{cat} of 800 s^{-1} compared to that of 290 s^{-1} for the His241 recombinant enzyme. The nitrophenyl riboside has been shown to be a good substrate for IU-nucleoside hydrolase but a poor substrate for *N*-ribohydrolases with more stringent substrate specificity or for purine nucleoside phosphorylase from calf spleen (Mazzella et al., 1996). Full activity of the mutant enzyme with nitrophenyl riboside establishes that the His241Ala enzyme folds to the catalytically active conformation. The low catalytic activity of the mutant with inosine establishes that His241 plays a significant role in hydrolysis of inosine but not for nitrophenyl riboside. The proposed role for the enzymatic acid−base pair is that the enzymatic proton donor, pK_a 9.1 and the enzymatic base, pK_a 7.1, function to protonate the hypoxanthine leaving group and to stabilize the developing charge on the ribosyl oxycarbonium ion, respectively (Parkin & Schramm, 1995). The pH profile for hydrolysis of nitrophenyl riboside by native IU-nucleoside hydrolase established that neither group is required for hydrolysis of this substrate but that the pK_a 7.1 base provides a 3-fold catalytic rate enhancement (Mazzella et al., 1996). The loss of the pK_a 9.1 group from the k_{cat} profile with nitrophenyl riboside implicates this group as the proton donor for the leaving group, since protonation is not required to activate the nitrophenylate ion as a leaving group.

Table 2: Kinetic Properties of Native, Recombinant, and His241Ala IU-Nucleoside Hydrolase[a]

	inosine			nitrophenyl riboside		
enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
native	31 ± 2	128 ± 30	2.5×10^5	220 ± 10	110 ± 10	2.0×10^6
recombinant	44 ± 3	150 ± 40	2.9×10^5	290 ± 6	66 ± 5	4.4×10^6
His241Ala	0.021 ± 0.005	530 ± 260	3.9×10^1	800 ± 20	840 ± 70	9.5×10^5

[a] Kinetic data for all of the enzymes were determined at pH 7.5 in 50 mM potassium phosphate and analyzed by fits to the Michaelis−Menten equation (Cleland, 1979).

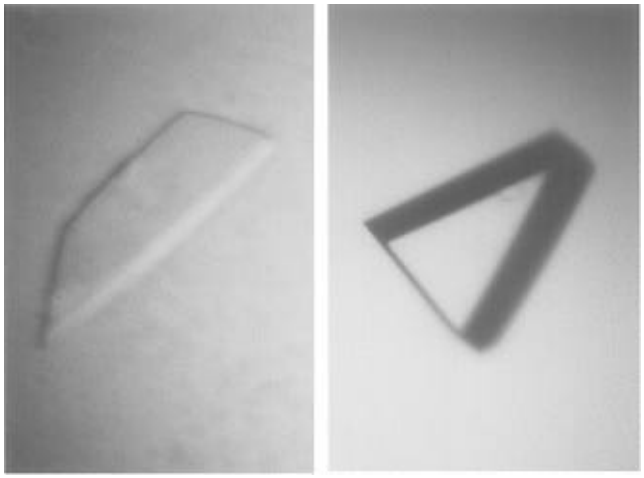

FIGURE 6: Crystals of native and recombinant IU-nucleoside hydrolase. In the left panel, IU-nucleoside hydrolase purified from *C. fasciculata* was crystallized using the conditions described in the text. In the right panel, the purified recombinant enzyme was crystallized in the same conditions. The largest crystal dimensions are approximately 0.5 mm for both left and right panels.

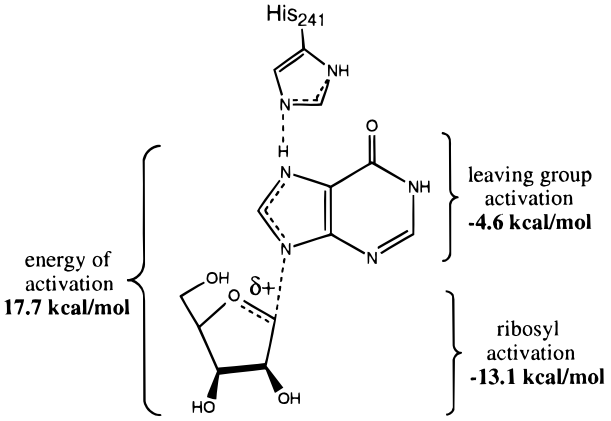

FIGURE 7: Proposed role of His241 in protonating the hypoxanthine leaving group at the transition state of IU-nucleoside hydrolase. The protonation is proposed from pH studies (Parkin & Schramm, 1995) and the crystallographic results in the following paper. The energetics are based on the relative substrate specificities on inosine and nitrophenyl riboside (Mazzella et al., 1996).

Assays of His241Ala IU-nucleoside hydrolase were conducted in variable concentrations of imidazole buffer, pH 7.0, to 0.5 M in attempts to restore catalytic activity with inosine as substrate. No effect of imidazole was observed. Thus, the His241Ala mutation does not permit catalytically competent interaction of exogenous imidazole with the catalytic site.

Energetics of Catalysis by Native and His241Ala IU-Nucleoside Hydrolases. Catalytic rate enhancement by purine N-ribohydrolases is estimated to be 6×10^{12} relative to solution solvolysis with the same conditions (DeWolf et al., 1979). The $\Delta\Delta G^{\ddagger}$ for enzymatic transition-state stabilization relative to that in solution can be estimated from the expression $\Delta\Delta G^{\ddagger} = -RT \ln$ (catalyzed/uncatalyzed). For native IU-nucleoside hydrolase, this calculation indicates a $\Delta\Delta G^{\ddagger}$ of 17.7 kcal mol^{-1} to be overcome by enzymatic interactions to achieve the transition state. The interaction of His241 with inosine at the transition state contributes a rate enhancement of 2100-fold, which represents −4.6 kcal/mol of the energy for catalysis. The remaining −13.1 kcal/mol therefore comes from interactions which distort the ribosyl toward the oxycarbonium ion transition state. The $\Delta\Delta G^{\ddagger}$ for hydrolysis of nitrophenyl riboside by IU-nucleoside hydrolase can be estimated to be 13.7 kcal mol^{-1} on the basis of the relative solvolytic and enzymatic rates.[2] The *p*-nitrophenylate anion is a good leaving group which does not require protonation. Contribution of −13.1 kcal/mol from the ribosyl-destabilizing interactions at the catalytic site is sufficient to give rapid solvolysis of the nitrophenyl riboside with or without interaction with His241. Although His241 is necessary for efficient catalysis of inosine, His241Ala *increases* k_{cat} for solvolysis of nitrophenyl riboside by a factor of 2.8 (Table 2). Ion pairing of the protonated His241 with the nitro group of the *p*-nitrophenylate anion or its bulk acting to block diffusion of the *p*-nitrophenylate from the catalytic site could be responsible for a decreased rate of product release and the observed decrease in k_{cat} with this substrate.

These results permit an energetically consistent assignment of the forces used to reach the transition state with native IU-nucleoside hydrolase. The major contribution of approximately −13 kcal/mol is from distortion of the ribosyl group to the oxycarbonium and charge stabilization of the oxycarbonium ion. The groups responsible for this destabilization appear to be a cluster of four aspartic acid residues (Degano et al., 1996; following paper). Activation of the hypoxanthine leaving group can be accomplished by interaction (possibly protonation) with the acidic form of His241 and accounts for the remaining −4.6 kcal/mol of activation energy. These interactions are summarized in Figure 7.

Summary. IU-nucleoside hydrolase from the trypanosome *C. fasciculata* is the first of the protozoan nucleoside hydrolases to be characterized by genetic and structural means. Similar proteins appear to exist in *E. coli*, yeast, and other protozoa but are unknown in mammals. The protein forms diffraction-quality crystals. The catalytic site requires His241 for the efficient hydrolysis of inosine but not for nitrophenyl riboside. The substrate specificity and kinetic constants are consistent with His241 acting as a proton donor to activate the hypoxanthine leaving group. Of

[2] The rate of solvolysis for nitrophenyl riboside at pH 7 is $10^{-7.5}$ s^{-1} compared to a k_{cat} of 239 s^{-1} for IU-nucleoside hydrolase (Mazzella et al., 1996). The enzymatic rate enhancement is therefore 8×10^9, corresponding to a $\Delta\Delta G^{\ddagger}$ for this substrate of 13.7 kcal/mol for catalysis by IU-nucleoside hydrolase.

the −17.7 kcal/mol contributed to transition state stabilization by the enzyme, −4.6 kcal/mol is from interactions with His241 and −13.1 kcal/mol is from enzymatic forces which distort the ribosyl toward the oxycarbonium.

ACKNOWLEDGMENT

The authors thank Drs. R. H. Angeletti, X.-Y. Tang, and E. Nieves in the Laboratory for Macromolecular Analysis for their generous assistance in peptide sequencing and electrospray mass spectrometry. We thank Dr. Buddy Ullman for advice on cloning from trypanosomes. The skilled manuscript preparation of Ms. Joan Byrne is appreciated.

REFERENCES

Alonso, G., Guevara, P., & Ramirez, J. L. (1992) *Mem. Inst. Oswaldo Cruz 87*, 517−523.
Boutellier, M., Horenstein, B. A., Semenyaka, A., Schramm, V. L., & Ganem, B. (1994) *Biochemistry 33*, 3994−4000.
Bouvier, J., & Stragier, P. (1991) *Nucleic Acids Res. 19*, 180.
Cleland, W. W. (1979) *Methods Enzymol. 63*, 103−138.
Degano, M., Gopaul, D. N., Scapin, G., Schramm, V. L., & Sacchettini, J. C. (1996) *Biochemistry 35*, 5971−5981.
DeWolf, W. E., Jr., Fullin, F. A., & Schramm, V. L. (1979) *J. Biol. Chem. 254*, 10868−10875.
Estupiñán, B., & Schramm, V. L. (1994) *J. Biol. Chem. 269*, 23068−23073.
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.
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.
Horenstein, B. A., Zabinski, R. F., & Schramm, V. L. (1993) *Tetrahedron Lett. 34*, 7213−7216.
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.
Koszalka, G. W., & Krenitsky, T. A. (1979) *J. Biol. Chem. 254*, 8185−8193.
Leung, H. B., & Schramm, V. L. (1984) *J. Biol. Chem. 259*, 6972−6978.
Leung, H. B., Kvalnes-Krick, K. L., Meyer, S. L., deRiel, J. K., & Schramm, V. L. (1989) *Biochemistry 28*, 8726−8733.
Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
Mazzella, L. J., Parkin, D. W., Tyler, P. W., 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. 259*, 5073−5077.
Muhich, M. L., Hughes, D. E., Simpson, A. M., & Simpson, L. (1987) *Nucleic Acids Res. 15*, 3141−3153.
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.
Rost, B., & Sander, C. (1993) *J. Mol. Biol. 232*, 584−599.
Rost, B., & Sander, C. (1994) *Proteins 19*, 55−72.

BI952998U