Ester Hydrolysis by dUTPase*  

Orsolya Barabási†, Veronika Pongrácz‡, Júlia Kovári†, Matthias Wilmanns‡, and Beáta G. Vértessy§‡  

From the †Institute of Enzymology, Biological Research Center, Hungarian Academy of Science, Budapest, Karolina út 29-31, H-1113, Hungary, the ‡Department of Theoretical Chemistry, Eötvös Loránd University, Budapest, H-1117 Hungary, and the §European Molecular Biology Laboratory, Hamburg Outstation, Hamburg, D-22603, Germany  

Received for publication, June 2, 2004, and in revised form, June 15, 2004. Published, JBC Papers in Press, June 17, 2004, DOI 10.1074/jbc.M406135200  

dUTPase is essential to keep uracil out of DNA. Crystal structures of substrate (dUTP and αβ-imino-dUTP) and product complexes of wild type and mutant dUTPases were determined to reveal how an enzyme responsible for DNA integrity functions. A kinetic analysis of wild type and mutant dUTPases was performed to obtain relevant mechanistic information in solution. Substrate hydrolysis is shown to be initiated via in-line nucleophile attack of a water molecule oriented by an activating conserved aspartate residue. Substrate binding in a catalytically competent conformation is achieved by (i) multiple interactions of the triphosphate moiety with catalytically-assisting Mg2+, (ii) a concerted motion of residues from three conserved enzyme motifs as compared with the apoenzyme, and (iii) an intricate hydrogen-bonding network that includes several water molecules in the active site. Results provide an understanding for the catalytic role of conserved residues in dUTPases.  

Due to the chemical reactivity of DNA, numerous mutagenic and carcinogenic modifications to mammalian genome occur daily (1). Preserving DNA integrity is of vital importance. The ubiquitous enzyme dUTPase performs a key role in preventing uracil incorporation into DNA by catalyzing the cleavage of dUTP into dUMP and pyrophosphate (2). This reaction contributes to thymidylate biosynthesis and strictly controls cellular dUTP/dTTP ratios (3).  

Targeting enzymes of de novo thymidylate biosynthesis by fluorouracil or methotrexate is a widely used approach in anticancer chemotherapy (4). These drugs perturb the cellular dUTP/dTTP pool resulting in synthesis of highly uracil-substituted DNA. Uracil-DNA transposes base-excision repair into a hyperactive cycle inducing DNA double-strand breaks and thymine-less cell death. This pathway preferentially kills cells actively synthesizing DNA, such as tumor and virus-infected cells. Moreover, in human cancer cell lines, dUTPase overexpression leads to development of fluorouracil resistance (5). Therefore, the clinical benefits of an anticancer therapy based on inducing thymine-less cell death may ultimately depend on the critical interplay among several enzymes involved in thymidylate metabolism. The recent finding of activation of dUTPase gene expression in p53 mutant tumor cells has reinforced the notion that this enzyme plays a significant role in tumor development and survival (6). The potential initiating role of dUTPase antagonism in thymine-less cell death has prompted investigations into the enzymatic mechanism.  

Detailed description of the catalytic mechanism is central to enzymology (7–9). Although several dUTPase crystal structures at atomic to moderate resolution have already been published (10–14), mechanistic issues remain obscure. Using water labeled with 18O, it was established that the oxygen of the substrate hydrolysis proceeds via a bridging network of residues, including the catalytic water molecule. Results provided a clear understanding of the catalytic role of conserved residues.  

In the present work, we addressed the unresolved mechanistic issues by three-dimensional structural investigations and kinetic/mutagenesis studies. We determined crystal structures of the wild type E. coli dUTPase:αβ-imino-dUTP:Mg2+, wild type dUTPase:dUMP, Asp90→Asn mutant dUTPase:dUTP/Mg2+ and Asp90→Asn mutant dUTPase:αβ-imino-dUTP:Mg2+ complexes (at 1.9, 1.6, 1.95, and 1.7 Å resolution) complexes to identify conformational changes of protein side chains and their interactions with the ligand. A kinetic analysis was performed for the reaction of αβ-imino-dUTP hydrolysis by wild type dUTPase, as well as for the reaction of dUTP hydrolysis by Asp90→Asn mutant dUTPase. Stepwise comparisons among these structures and the apoenzyme structure (11) together with the kinetic results revealed essential interactions required for catalysis and allowed novel insights into the mechanism.
Experimental Procedures

Chemicals, Enzyme Preparation, and Purification—α,β-imino-dUTP was obtained from Jena Bioscience. α,β-imino-dUTP was prepared by enzymatic phosphorylation of dUMP and α,β-imino-dUTP (18) and was shown to be over 98% pure by ion exchange chromatography (Fig. 1A). Other chemicals were of proanalysis quality from either Merck or Sigma. E. coli dUTPase and 1 enzyme–built with ease.

Any model bias with respect to the ligand position. The asymmetric unit search model for all the four presently determined structures to exclude

A nitrogen stream. Three data sets were collected at various synchrotron amounts of ligand in reservoir solution, then flash-frozen in a 100 K was followed in aliquots taken at different time points to be analyzed either on a Mono Q anion exchange column or by thin layer chromatography (17).

Protein Crystallization and Data Collection—Crystals were obtained at 20 °C by hanging-drop vapor diffusion. The protein/ligand solution prepared for co-crystallization contained 3 mg/ml enzyme, 1.25 mM α,β-imino-dUTP, or 5 mM dUTP, or 1.25 mM dUMP, and 10 mM MgCl₂ in 10 mM Tris/HCl buffer, pH 7.0, and 50 mM NaCl. dUTPase:ligand: α,β-imino-dUTP was mixed with equal volume of reservoir solution (0.1 M Tris/HCl buffer, pH 7.8, 5 mM MgCl₂, also containing 400 mM sodium acetate, at room temperature. The decrease of α,β-imino-dUTP levels and production of dUMP was followed in aliquots taken at different time points to be analyzed either on a Mono Q anion exchange column or by thin layer chromatography (17).

Structure Determination and Refinement—Structures of wild type E. coli dUTPase and the enzyme–substrate complexes with lifetimes allowing crystallographic characterization were considered, the following potentials.

First, the substrate dUTP was replaced by the substrate analogue α,β-imino-dUTP. Imino analogues of nucleotide phosphates are commonly used as isosteric mimics of the natural substrates (31, 32) both with respect to bond distances (1.63 and 1.68 Å, for P–O–P and P–N–P moieties, respectively) and bond angles (128.7° and 127.2°, for P–O–P and P–N–P, respectively) (33). The lower electron negativity of the nitrogen atom results in less reactivity of the phosphate ester imino analogues. Nevertheless, several enzymes (e.g. small GTPase p21 (34), sarcoplasm reticulum ATPase (35), and alkaline phosphatase (36)) were reported to catalyze imino-linkage cleavage, arguing for some reactivity associated with the P–N–P moiety.

The Experimental Approach—Previous dUTPase complex structures identified the protein fold and the overall characteristics of active site architecture; however, they failed to provide an understanding of the catalytic mechanism (10, 12–14). This failure was probably due to the fact that these structures represented dead-end enzyme-inhibitor complexes. To produce the relevant enzyme-substrate complexes with lifetimes allowing crystallographic characterization, we considered the following potentials.

The decrease of α,β-imino-dUTP levels and production of dUMP was followed in aliquots taken at different time points to be analyzed either on a Mono Q anion exchange column or by thin layer chromatography (17).

Protein Crystallization and Data Collection—Crystals were obtained at 20 °C by hanging-drop vapor diffusion. The protein/ligand solution prepared for co-crystallization contained 3 mg/ml enzyme, 1.25 mM α,β-imino-dUTP, or 5 mM dUTP, or 1.25 mM dUMP, and 10 mM MgCl₂ in 10 mM Tris/HCl buffer, pH 7.0, and 50 mM NaCl. dUTPase:ligand: α,β-imino-dUTP was mixed with equal volume of reservoir solution (0.1 M Tris/HCl buffer, pH 7.8, 5 mM MgCl₂, also containing 400 mM sodium acetate, at room temperature. The decrease of α,β-imino-dUTP levels and production of dUMP was followed in aliquots taken at different time points to be analyzed either on a Mono Q anion exchange column or by thin layer chromatography (17).

Catalytic competence of the enzyme in the crystal phase was checked by competition experiments with the native substrate dUTP. Crystals of wild type dUTPase containing α,β-imino-dUTP were washed with reservoir solution and dUTP was added at 10 mM concentration. Hydrolysis was followed by discontinuous thin layer chromatography activity test (24).

Structure Determination and Refinement—Structures of wild type and Asp⁹⁰ → Asn mutant dUTPases, in complexes with α,β-imino-dUTP:Mg²⁺, dUTP:Mg²⁺, and dUMP, were solved by molecular replacement (MOLREP (25)). The apo-dUTPase structure (11) was used as search model for all the four presently determined structures to exclude any model bias with respect to the ligand position. The asymmetric unit contained one monomer. In each case, the resultant initial maps were of exceptional quality, allowing residues 1-136 and entire ligands to be built with ease.

Generation of monomer libraries for the ligands (α,β-imino-dUTP, dUTP) and refinement were carried out using CCP4i and Refmac5 (22). Positional and B-factor refinement rounds were altered with manual rebuilding steps using the graphics program O (26), and ARP/WARP (27) was used for water building. Residues belonging to the C-terminal M2 helix (residues 137–152) are hardly visible even in the final maps and are therefore mostly omitted from the model. A summary of the crystallographic data collection and refinement statistics is given in Table II.

Coordinates and structure factor data have been deposited in the Protein Data Bank with accession codes 1RN8, 1RNJ, 1SYL, and 1SEH, for the wild-type enzyme-α,β-imino-dUTP complex, the Asp⁹⁰ → Asn enzyme-α,β-imino-dUTP complex, the Asp⁹⁰ → Asn enzyme-dUTP complex, and the wild type enzyme-dUTP complex, respectively.

The figures were produced using Molscript v2.1.2 (28), Raster3D v2.7b (29), and Bobscript v2.6b (30).

Simulated annealed omit electron density maps (sigma-weighted 2Fo − Fc, maps) were calculated using CNS v1.1, omitting the entity in question and a 3.5-Å region surrounding it. The figures are restricted to show only the entity in question.

RESULTS

The Experimental Approach—Previous dUTPase complex structures identified the protein fold and the overall characteristics of active site architecture; however, they failed to provide an understanding of the catalytic mechanism (10, 12–14). This failure was probably due to the fact that these structures represented dead-end enzyme-inhibitor complexes. To produce the relevant enzyme-substrate complexes with lifetimes allowing crystallographic characterization, we considered the following potentials.

First, the substrate dUTP was replaced by the substrate analogue α,β-imino-dUTP. Imino analogues of nucleotide phosphates are commonly used as isosteric mimics of the natural substrates (31, 32) both with respect to bond distances (1.63 and 1.68 Å, for P–O–P and P–N–P moieties, respectively) and bond angles (128.7° and 127.2°, for P–O–P and P–N–P, respectively) (33). The lower electron negativity of the nitrogen atom results in less reactivity of the phosphate ester imino analogues. Nevertheless, several enzymes (e.g. small GTPase p21 (34), sarcoplasm reticulum ATPase (35), and alkaline phosphatase (36)) were reported to catalyze imino-linkage cleavage, arguing for some reactivity associated with the P–N–P moiety.

Fig. 1A demonstrates that dUTPase is also capable of hydrolyzing the imino-linkage of α,β-imino-dUTP, while no detectable hydrolysis of the β–γ bond occurs, i.e. the exclusive specificity of dUTPases for α–β bond cleavage is retained in the reaction with the imino analogue as well. Table I summarizes kinetic and ligand binding data for the dUTPase-catalyzed dUTP and α,β-imino-dUTP hydrolysis. The very low kcat value in the α,β-imino-dUTP hydrolysis reaction rendered Kcat determination unreliable; therefore the previously determined Kcat value is presented for comparison (37). Assuming rapid equilibrium reactions, as suggested by an earlier kinetic analysis (15), Kn and Kd values are expected to be the same. The very close agreement in the nucleotide interaction data (Kd, Kn, Kd, cf. Table I) for dUTP and the analogue α,β-imino-dUTP strongly suggests highly similar accommodation patterns for these two hydrolysable nucleotides in the dUTPase active site, although the rate of hydrolysis differs by almost 6 orders of magnitude. Based on these kinetic and ligand binding data, both the enzyme-dUTP and the enzyme-α,β-imino-dUTP complex will be termed in the present work as enzyme-substrate complexes (Fig. 1B).

Second, the dUTPase catalytic co-factor Mg²⁺ ion (15, 16), not yet localized in any previously reported crystal structure (10–14, 38, 39), was also included during complex formation. Mg²⁺ binding in the enzyme-α,β-imino-dUTP/dUTP complexes was promoted by novel crystallization conditions at neutral or slightly basic pH values, since deprotonation of side chains has a well known positive effect on metal ion coordination capability. Earlier crystallization solvents contained either acidic buffers (11, 12, 38) or high concentrations of metal-chelating citrate (13, 39).

Third, a site-directed mutation was designed to exchange the strictly conserved aspartate side chain (Asp⁹⁰) within the active site into an asparagine residue. This mutation was expected to decrease the reaction rate and allow for a structural analysis of the enzyme-dUTP complex in the presence of the co-factor Mg²⁺ ion.

Structure of Enzyme-Substrate Complex—dUTPases show exquisite substrate specificity with respect to base, sugar, and phosphate chain moieties of dUTP that is essential to prevent unintended hydrolysis of other high energy-containing nucleotides. Most dUTPases are homotrimers (Fig. 2A) with five conserved sequence motifs in each monomer (Fig. 2B) (10, 12–14). Consensus sequences for conserved dUTPase motifs
are: Motif I, Ala-Gly-(Pho)-Asp-Leu; Motif II, (Pro/Gly)-(Arg/Lys)-Ser; Motif III, Gly-(Pho)2-Asp-(Nnn)2-Tyr-(Nnn)-Gly; Motif IV, Arg-(Pho)-Ala-Gln; Motif V, Arg-Gly-(Nnn)2-Gly-Phe-Gly-(Ser/His)-(Thr/Ser)-Gly (amino acids in three-letter codes: Pho, hydrophobic residue; Nnn, any residue) (10, 40). Overall fold and active site architecture are notably well conserved.

**Fig. 1. Reactions catalyzed by dUTPase.** A, experimental observation of α,β-imino-dUTP hydrolysis. Reaction mixtures (cf. “Experimental Procedures”) were set up in the absence (thick dashed trace) or in the presence (thick solid trace) of E. coli dUTPase and were incubated for 2 weeks at room temperature. Chromatograms represent the nucleotide composition; arrows indicate the corresponding elution positions of dUMP and α,β-imino-dUTP standards. The thin line represents the salt gradient applied in the chromatography. B, theoretical reaction schemes of dUTP (X = O) or α,β-imino-dUTP (X = N) hydrolysis. The catalytic water molecule initiating nucleophile attack and the Mg\(^{2+}\) ion coordinated to the triphosphate chain are shown as determined in the crystal structure (D). In the reaction products, the Mg\(^{2+}\)-pyrophosphate interaction is hypothetic, the hypothesis being based upon the apparent absence of the Mg\(^{2+}\) from the dUTPase:dUMP structure. C, D, and E, active-site close-ups of apoenzyme, dUTPase:α,β-imino-dUTP:Mg\(^{2+}\), and dUTPase:dUMP structures, respectively. In the subunit-color-coded ribbon model of the protein, the side chain and/or main chain atoms of some conserved residues (Tyr\(^{93}\), Asp\(^{90}\), Leu\(^{88}\) from monomer A, Ala\(^{29}\), Asp\(^{32}\), Arg\(^{71}\), Ser\(^{72}\), Gly\(^{73}\), and Gln\(^{119}\) from monomer B) important in active site architecture are also shown, together with some significant water molecules. For denominations of these residues and waters, omitted from this figure to avoid small print details, please refer to Fig. 3A that presents the active site in the same orientation. The ligands and the non-carbon protein atoms are in atom coloring bonds model (carbon, dark gray; oxygen, red; phosphorus, orange; nitrogen, blue; magnesium, purple). The apoenzyme structure was taken from the literature (11).
among dUTPases from retroviruses and bacteria to man (10, 12–14).

To study the first step of the reaction, the structure of the dUTPase: α,β-imino-dUTP:Mg2⁺ complex (Figs. 1D, 2A, 2B, 2C, and 4A and Table II) was determined at 1.9-Å resolution. In agreement with previous studies (10, 12–14), each α,β-pleated subunit forms jellyroll topology and contributes the C-terminal α-strand to the neighboring subunit. The three active sites of the homotrimer are located in clefts between neighboring monomers and recruit conserved motifs from different subunits in a 3-fold symmetric pattern (Fig. 2A). A β-hairpin formed by conserved Motif III from monomer A accommodates the uracil and sugar rings. A conserved tyrosine (Tyr93) stacks to the 2'-deoxyribose ring and sterically excludes ribonucleotides. Phosphate chain recognition is provided by conserved Motifs I, II, and IV of monomer B (Figs. 1D, 4A and 4B, and 3). Therefore, active site architecture ultimately depends on correct oligomerization of the homotrimer. To our present knowledge, this organization of ligand binding sites is unique among proteins.

Table I
Kinetic and ligand binding data for wild type and Asp90 → Asn mutant E. coli dUTPase

| Nucleotide          | Wild type | Asp90 → Asn mutant |
|---------------------|-----------|--------------------|
|                     | kcat      | Km                 | kcat/Km | kcat      | Km   | kcat/Km |
| dUTP                | 11 s⁻¹    | 0.5 μM             | 1.4 × 10⁷ | 2.55 × 10⁻⁴ | 0.45 μM | 850 |
| α,β-imino-dUTP      | 2 × 10⁻⁵  | 1.0 μM⁺            | 20⁶     | ND        | ND   | ND     |

* Kd from Ref. 37.  
* kcat/Kd.

Fig. 2. A, overall structure of dUTPase. Enzyme-substrate (E. coli dUTPase: α,β-imino-dUTP:Mg²⁺) complex. The trimer is shown with ribbon model of color-coded (code retained throughout the present study, where applicable) subunits and bonds model of the nucleotide ligand molecule in the three active sites. Mg²⁺-ions are represented as purple balls. B, Sequence alignments of dUTPases. Conserved residues within the five (I-V) dUTPase motifs are in bold white on black background. Arrowhead points at the conserved aspartate in Motif III, Asp90 in the E. coli sequence, which was selected for the Asn mutation.
TABLE II
Crystallographic data collection and refinement statistics

|                  | WT-dUTPase αβ-imino-dUTP:Mg\(^{2+}\) | D90N mutant dUTPase αβ-imino-dUTP:Mg\(^{2+}\) | D90N mutant dUTPase:dUTP:Mg\(^{2+}\) | WT-dUTPase:dUMP |
|------------------|-----------------------------------|-----------------------------------------------|-----------------------------------|----------------|
| Space group      | P6,22                             | P6,22                                         | P6,22                             | P6,22          |
| Cell parameters α, c (Å) | 74.6, 99.6                       | 74.9, 99.6                                   | 75.3, 98.8                       | 75.3, 100.5    |
| Data collection  |                                   |                                               |                                   |                |
| Wavelength (Å)   | 0.811                             | 0.9392                                        | 0.8414                           | 0.8040         |
| Resolution (Å)   | 27.1–1.9                          | 22.0–1.7                                      | 20.0–1.91                        | 20.0–1.5       |
| (1.98–1.93)      | (1.74–1.70)                       | (2.03–1.91)                                  | (1.60–1.47)                      |                |
| Measured reflections | 128,620 (6,183)                  | 372,822 (26,388)                             | 38,679 (6,077)                  | 930,805 (129,002) |
| Unique reflections | 12,804 (819)                    | 18,346 (1,288)                               | 13,413 (2,085)                  | 29,251 (6,421) |
| Completeness (%) | 99.2 (99.2)                       | 98.2 (98.2)                                  | 99.3 (98.5)                     | 99.9 (100.0) |
| I/σ(I)           | 10.3 (2.8)                        | 8.4 (2.0)                                    | 15.0 (2.5)                      | 27.6 (4.9)     |
| R\(_{\text{sym}}\) (%) | 6.1 (26.8)                      | 5.9 (38.0)                                   | 6.3 (48.4)                      | 8.3 (73.0)     |
| Refinement statistics |                                 |                                               |                                   |                |
| Resolution (Å)   | 20.0–1.93                         | 20.0–1.70                                    | 20.0–1.95                       | 20.0–1.47      |
| Nonhydrogen atoms | 1317                             | 1353                                         | 1250                            | 1298           |
| Water molecules  | 191                               | 221                                          | 167                             | 194            |
| Data-parameter restraint | 12,892:5,278:16,976 | 18,811:5,422:17,195                     | 13,200:5,010:16,328             | 29,279:5,202:16,906 |
| r.m.s. deviation bonds (Å) | 0.016                            | 0.019                                        | 0.020                           | 0.021          |
| r.m.s. deviation angles (°) | 1.60                             | 1.54                                         | 1.69                            | 1.38           |
| R\(_{\text{work}}\) (%) | 13.7                             | 15.7                                         | 15.8                            | 15.6           |
| R\(_{\text{free}}\) (%) | 18.0                             | 18.7                                         | 18.9                            | 18.4           |

\( R_{\text{sym}} = \frac{\sum |I_j - \langle I \rangle|}{\sum I} \), where \( I \) is the recorded intensity of the \( j \)th reflection and \( \langle I \rangle \) is the average intensity over multiple recordings.

\( R_{\text{free}} \) values are calculated for a randomly selected 5% of the data that was excluded from the refinement.

Comparison of apoenzyme and enzyme-substrate complex shows only minor conformational shifts in the active site, and most of these are involved in residues of subunit B, responsible for phosphate accommodation. The movement of Arg\(^{116}\) and Arg\(^{23}\) by 1.23 and 0.88 Å, respectively, provide phosphate chain interactions, while Gln\(^{119}\) adopts an altered conformation for proper hydrogen-bonding geometry. The catalytic site in the apoenzyme already presents side chain conformations adequate for binding of the uracil and deoxyribose moieties, alleviating the need for significant changes induced by the incoming dUTP (Figs. 1, C and D, 3D).

Identification of the Nucleophile Responsible for Attack at the α-Phosphorus—The present dUTPase-substrate complex structure allows identification of the attacking nucleophile. A search along the axis of the α-phosphorus-imino nitrogen bond located a single possible candidate for in-line nucleophile attack: a water oxygen (termed W\(_{\text{cat}}\), numbered as W5 in the 1RN8 coordinate file, red arrow in Figs. 1D and 3A and B, cf. also Fig 4A) coordinated to Asp\(^{90}\) in monomer A (previously suggested to be a catalytic residue (12)). No other protein or water atoms were found within 4.0 Å to the α-phosphorus, emphasizing the special role of the proximal water (W\(_{\text{cat}}\)). A simulated annealed omit map, calculated with omission of entities in the 3.5 Å radius region around this water, was created for its clear localization (Fig. 3A). Both carboxyl oxygens of Asp\(^{90}\) may participate in hydrogen-bonding to W\(_{\text{cat}}\). AspO\(_{\text{cat}}\) is also within H-bonding distance to the deoxyribose 3′-OH group, while AspO\(_{\text{cat}}\) is H-bonded to monomer B Ala\(_{27}\) main chain NH and another water (W4) that contacts monomer B Gln119 through W1 (Figs. 3A and 4A). The monomer B Ala\(_{27}\) main chain H-bond may play a role in orienting the Asp\(^{90}\) side chain. No other side chain atoms were closer than 4.0 Å to this putative catalytic water. However, another protein atom (Leu88 main chain carbonyl O) is also within H-bonding distance to W\(_{\text{cat}}\) as discussed below. This arrangement suggested that replacing only one side chain oxygen of Asp\(^{90}\) with an amino group might drastically affect enzyme function. The Asp\(^{90}\) → Asn mutant was constructed and proved to be largely compromised in catalytic efficiency (Table I). k\(_{\text{cat}}\)/K\(_{\text{m}}\) for the mutant was determined to be 850 M\(^{-1}\) s\(^{-1}\) as compared with 3.5 × 10\(^7\) M\(^{-1}\) s\(^{-1}\) for wild type, a decrease of almost 5 orders of magnitude. Importantly, K\(_{\text{m}}\) was the same for both mutant and wild type enzymes (0.45 versus 0.5 μM) (Table I), indicating that substrate accommodation is the same in both species. The crystal structure of the Asp\(^{90}\) → Asn mutant in complex with Mg\(^{2+}\) and α,β-imino-dUTP was solved at 1.7Å resolution and is shown as superimposed on the wild type complex structure (Fig. 3B, cf. also Fig 3F, and Table II). The superimposed structures are practically indistinguishable (r.m.s.\(^3\) for fitting of all atoms is 0.12), except for the catalytic water electron density (Fig. 3A, B, E, and F) that is missing from the mutant complex structure. A straightforward interpretation of this result is that replacing AspO\(_{\text{cat}}\) with an NH\(_2\) group with altered H-bonding characteristics adversely interferes with W\(_{\text{cat}}\) coordination in the Asp\(^{90}\) → Asn mutant. A rotation of AsnO\(_{\text{cat}}\) (the corresponding atom for AspO\(_{\text{cat}}\)) into the AspO\(_{\text{cat}}\) position could theoretically provide proper coordination for W\(_{\text{cat}}\) but this rotation is impeded by the AsnO\(_{\text{cat}}\)-monomer B Ala\(_{27}\) main chain H-bond, also present with AspO\(_{\text{cat}}\) in the wild type complex.

\(^3\) The abbreviation used is: r.m.s., root mean square.
The very low enzymatic activity of the Asp<sup>90</sup>→Asn mutant rendered it possible to determine the crystal structure of the mutant enzyme:dUTP:Mg<sup>2+</sup> complex. In addition to the catalytic water, Mg<sup>2+</sup>-coordinating waters, W1, W2, W4, W15, and W21, also participate in the primary hydrogen-bonding interactions. B, superimposed structures of wild type (dark tones) and Asp<sup>90</sup>→Asn mutant (light tones) dUTPase:α,β-imino-dUTP:Mg<sup>2+</sup> complexes. Note that the only remarkable difference between the superimposed structures is the disappearance of W<sub>cat</sub> from the mutant complex. Atomic color code: carbon, dark/light gray; oxygen, dark/light red (pink); phosphorus, dark/light orange (yellow); nitrogen, dark/light blue; magnesium, dark/light purple. C, superimposed structures of Asp<sup>90</sup>→Asn mutant dUTPase: dUTP:Mg<sup>2+</sup> (dark tones) and Asp<sup>90</sup>→Asn mutant dUTPase:α,β-imino-dUTP:Mg<sup>2+</sup> (light tones) complexes. Note the close identity in the positions of the nucleotide ligands. D, apoenzyme retains a water closely corresponding to the W<sub>cat</sub> position. 3-Fold superimposition of the apoenzyme (green carbons and water, otherwise standard atom coloring), enzyme-substrate (dark tones), and enzyme-product (light tones) structures. Note the position of the catalytic water from the apoenzyme to the enzyme-substrate and enzyme-product complexes. E, F, and G, simulated annealed omit electron density maps for the substrates in wild type *E. coli* dUTPase:α,β-imino-dUTP:Mg<sup>2+</sup>, the Asp<sup>90</sup>→Asn *E. coli* dUTPase:α,β-imino-dUTP:Mg<sup>2+</sup>, and the Asp<sup>90</sup>→Asn *E. coli* dUTPase:dUTP:Mg<sup>2+</sup> structures, respectively. Maps are restricted to show the nucleotide ligand, the Mg<sup>2+</sup>, the three water molecules coordinating to the metal ion, as well as the catalytic water, if present.

The very low enzymatic activity of the Asp<sup>90</sup>→Asn mutant rendered it possible to determine the crystal structure of the mutant enzyme:dUTP:Mg<sup>2+</sup> complex (Figs. 3, C and G, Table II). In this structure, substrate and Mg<sup>2+</sup> accommodation patterns are highly similar to those realized in the enzyme:α,β-imino-dUTP:Mg<sup>2+</sup> complex (Fig. 3, C, F, and G). All atoms of the imino analogue and the physiological substrate, including the α-phosphate group where nucleophilic attack occurs, are well superimposable (r.m.s. is 0.179).

These structures, together with the kinetic data, clearly demonstrate that the water molecule termed W<sub>cat</sub> provides the attacking nucleophile oxygen. Although the structures that identified W<sub>cat</sub> have been determined for the α,β-imino-dUTP complexes of wild type and Asp<sup>90</sup>→Asn mutant dUTPsases, the catalytic incompetence of the Asp<sup>90</sup>→Asn mutant was proved in the physiological reaction (dUTP hydrolysis). Therefore, the Asp<sup>90</sup> residue plays a determinant role in dUTP hydrolysis, too. Two roles have been associated with this residue: (i) coordination of deoxyribose 3′OH (Figs. 3A and 4A, cf. also Refs. 10, 12–14) and (ii) H-bonding to W<sub>cat</sub> (Figs. 3A and 4A). The Asp<sup>90</sup>→Asn (i.e. Asp<sup>90</sup>Oδ2→Asn<sup>−</sup>Nδ2) mutant is not compromised in the first role (cf. equal accommodation of dUTP as...
evidenced by $K_m$ values in Table I and three-dimensional structures in Fig. 3, B, E, and F). Consequently the attenuated catalytic efficiency of the mutant enzyme is most probably due to lack of proper coordination to the catalytic water molecule in the physiological reaction. In addition, phosphate chain conformations in the dUTPase:α,β-imino-dUTP:Mg$^{2+}$ and dUTPase: dUTP:Mg$^{2+}$ complexes are identical (Fig. 3C), indicating that the nucleophile attack mechanism should be similar in both cases.

Interestingly, the catalytic water molecule seems to be present in the apoenzyme structure (1EUW (11)), as well (W334 at a distance of 0.86 Å to the Wcat position, cf. also Fig 3D). It is coordinated to Asp$^{90}$O$^{62}$, as well as to the Leu$^{98}$ main chain carbonyl oxygen, just as it is found in the presently determined enzyme:α,β-imino-dUTP:Mg$^{2+}$ complex (Fig. 3D). In the latter complex, Wcat clearly adopts a closely co-linear location to the scissile bond to carry out nucleophilic attack on the α-phosphate. In the enzyme-product complex structure, as detailed below (Table II, Figs. 1E, 3D, and 4B), the Wcat proximal water (W49) has again the same coordination pattern. In summary, Wcat (or the Wcat proximal water) coordinates to the same protein atom ligands (Asp$^{90}$O$^{62}$ and Leu$^{98}$ main chain O) in all these structures. We propose that these two protein ligands, with the possible participation of other water molecules, create a binding site for Wcat that is, in fact, one of the substrates required for the dUTP hydrolysis reaction (dUTP + H$_2$O = dUMP + PP$_i$, cf. Fig 1B). This binding site is already available in the absence of the nucleotide substrate. After completion of the catalytic reaction, products (dUMP and PP$_i$) are expected to be exchanged for substrates (dUTP and H$_2$O). Following the two-substrate analogy, the presence of Wcat in the enzyme-product complex may be interpreted as recruitment of a solvent water molecule in such an exchange reaction.

Completion of the Catalytic Cycle—The pyrophosphate group

FIG. 4. Interaction mapping in enzyme-substrate (A), and enzyme-product complexes (B). Interactions are shown only for the phosphate chain moiety of the ligand. Due to the close similarity of the nucleotide interactions in the three enzyme-substrate complexes determined in the present study (cf. Fig. 3 and Table I), the map was selected to show the actual distances as found in the wild type dUTPase:α,β-imino-dUTP:Mg$^{2+}$ (X = N) complex where Wcat is also present. In the Asp$^{90}$ → Asn mutant dUTPase:α,β-imino-dUTP:Mg$^{2+}$ (X = N) and Asp$^{90}$ → Asn mutant dUTPase:dUTP: Mg$^{2+}$ (X = O) complex, the only significant differences are that (i) Wcat is absent and Asp$^{90}$O$^{62}$ becomes AsnN$^{62}$ and (ii) in the Asp$^{90}$ → Asn mutant dUTPase:dUTP: Mg$^{2+}$ (X = O) complex, the X-Ser$^{72}$O$_Y$ interaction is absent. Changes in all other distances are within ±0.2 Å.
at the entrance of the active site is expected to exit freely as it can easily make numerous contacts with the bulk solvent. The metal ion may facilitate pyrophosphatase discharge since it is more likely to bind to the pyrophosphate (charge minus 4) as compared with dUMP (charge minus 2 at physiological pH). Accordingly, no electron density could be assigned to Mg$^{2+}$ in the enzyme-product structure, although the dUTPase-dUMP complex has been crystallized in the presence of the metal ion (Table II and Figs. 3D and 4B). The C-terminal Motif V from monomer C, mostly disordered in our present and also in previously reported structures, may also participate in pyrophosphatase escape via charge stabilization with its strictly conserved Arg$^{141}$ side chain. Preferential discharge of the pyrophosphatase moiety over the dUMP is also strengthened by x-ray crystallographic investigations carried out on dUTPase crystals soaked into substrate containing solution that retained dUMP, but not the pyrophosphatase moiety, in the enzyme active site (14).

Figs. 3D and 4, A and B, indicate that the reaction product dUMP retains most of the binding interactions found with the substrate with respect to the sugar and base moieties. However, some side chains providing interactions with the triphosphate moiety (e.g. monomer B side chains Asp$^{52}$, Arg$^{71}$, Ser$^{72}$, and Gln$^{119}$) are shifted back to their positions occupied in the apoenzyme (Fig. 3D). This shift may contribute to the significant decrease of dUMP binding affinity to dUTPase, when compared with dUDP, dUTP, and a,b-imino-dUTP (15, 37). Less tight binding of dUMP as compared with dUTP facilitates exchange of product with substrate in the active site whenever substrate is available in the bulk phase.

**DISCUSSION**

The side-by-side three-dimensional structural analysis of the apoenzyme, as well as substrate, and product complexes delineated atomic interactions crucial to the catalytic mechanism of dUTPase. The conclusions were based on structures and kinetic data determined with the imino substrate analogue, as well as the physiological substrate and the physiological product. Several results indicate that the reaction pathway may well be of similar character with both dUTP and a,b-imino-dUTP. First, a water molecule in the immediate vicinity to the presently determined W$_{cat}$ is localized already in the apoenzyme structure and is also present in the enzyme-dUMP complex (Fig. 3D). Second, the binding position of the physiological substrate dUTP, most importantly the entire phosphate chain conformation, is equivalent to that of the imino analogue (Fig. 3, C, F, and G). Third, the Asp$^{90}$ $\rightarrow$ Asn mutation was shown to adversely affect W$_{cat}$ coordination and, parallel to this, to inactivate the enzyme in the physiological reaction. It is therefore reasonable to assume that (i) a water molecule coordinated to Asp$^{90}$ and Leu$^{96}$ is present with high probability in the active site of the wild type enzyme when dUTP is also bound, and (ii) this water molecule is poised for in-line attack, well within 4.0 Å to the $\alpha$-phosphorus. Given such an arrangement, it can be assumed that this specific water will act as the nucleophile in dUTP hydrolysis.

**The Metal Ion Site**—The localization of the site for the catalysis-assisting Mg$^{2+}$ provided significant novel insights, as well. Its important role in providing catalytically competent accommodation of the substrate in the dUTPase active site, as judged by $k_{cat}$ and $k_m$ determinations in the presence and absence of the metal ion, was well established in a number of earlier studies on dUTPases from diverse sources (15–17, 24). It was also well known that divalent metal ions usually contribute to enzymatic reactions of nucleotides by coordinating the otherwise quite flexible phosphate chain. Consequently, absence of the active site Mg$^{2+}$ from the earlier determined crystal structures prevented the definition of the catalytically competent phosphate chain conformation, thereby rendering the identification of the nucleophile attacker impossible. The data on the dUTPase:dUDP: Sr$^{2+}$ complex were a significant step forward in this problem (10). However, due to the diminished catalytic competence of the Sr$^{2+}$ substitution as well as the altered coordination geometry of the Sr$^{2+}$ (coordination number 8) as compared with Mg$^{2+}$ (coordination number 6), the character of the nucleophile attack could not be determined beforehand. In fact, in the dUTPase:dUDP: Sr$^{2+}$ complex, the critical position of the $\alpha$-phosphorus is 3.2 Å away from its competent site as determined in our present dUTPase:dUTP:Mg$^{2+}$ and Mg$^{2+}:\alpha$-b-imino-dUTP:dUTPase structures. The coordination of Mg$^{2+}$ to all the three phosphate groups is not very common in other nucleotide-protein structures where the metal ion is frequently seen as coordinating to only two phosphate groups (41–44). In the dUTPase structures, the triple phosphate coordination of Mg$^{2+}$ contributes to an increasingly compact phosphate chain conformation presumably optimal for catalysis.

**Role of Conserved Residues**—The enzymatic reaction is facilitated by non-covalent bonding interactions and involves residues from different monomers, as well as several water molecules bound within the active sites. All residues proposed to interact with ligand and W$_{cat}$ are strictly conserved among dUTPase sequences. The contacts and role of some of these residues, like those of the $\beta$-hairpin, Ile$^{69}$, Asp$^{90}$, Tyr$^{92}$, Ser$^{72}$, Arg$^{71}$, Arg$^{116}$, and Gln$^{119}$, were already suggested based on previously determined crystal structures of E. coli, equine infectious anaemia virus (EIAV), feline immunodeficiency virus (FIV), and Homo sapiens dUTPases (10, 12–14). However, as previous studies failed to define a catalytically competent substrate conformation, the relevant roles of conserved residues need to be reinterpreted in the light of our present data.

In general, the conclusions drawn from our structures concerning the role of Gly$^{73}$ main chain nitrogen and Arg$^{71}$ guanidino group in coordination and charge distribution of the phosphate chain, especially the $\beta$- and $\gamma$-phosphates correspond well with suggestions supported by previously determined structures of dUTPases from various sources (10, 12–14). However, the interaction between the phosphate chain and Arg$^{71}$ became water-mediated in our structures. This difference might be ascribed to the effective charge stabilization provided by the Mg$^{2+}$ in our structures. Due to the presence of Mg$^{2+}$, the triphosphate chain adopts a more compact conformation that increases its distance from the Arg$^{71}$ guanidino group. The weaker charge stabilization effect of the Arg$^{71}$ side chain is probably compensated for by the Mg$^{2+}$ ion.

The compact character of the phosphate chain in the present structures also shifts the interaction between the Ser$^{72}$ main chain NH and an oxygen in the $\beta$-phosphate group to another oxygen in the $\alpha$-phosphate group. This shift may contribute to activation of the $\alpha$-phosphorus for the nucleophile attack and might stabilize the developing charge in the transition state.

The present data provide detailed characterization of role of Asp$^{92}$ as well. In the previous retroviral dUTPase:dUDP: Sr$^{2+}$ complex structure (10), one of the carboxyl oxygens of this conserved residue (Asp$^{90}$ in the retroviral sequence) was within H-bonding distance to one of the eight coordinating ligands of the Sr$^{2+}$ ion. The large size of Sr$^{2+}$, its altered coordination geometry, and poor catalytic competence, however, attenuated the relevance of the Mg$^{2+}$ $\rightarrow$ Sr$^{2+}$ substitution for mechanistic conclusions. In the present structures, both side chain oxygens of Asp$^{92}$ take part in coordination of two water molecules in the coordination sphere of the physiological metal ion that has six ligands.

**The Catalytic Water**—We presented a description of the catalytic pathway via identification of the nucleophile agent and
characterization of interactions responsible for building up the required arrangement and initiating the reaction. The key residue here, Asp<sup>90</sup>, was previously shown to contribute to substrate binding via H-bonding to the OG group of the ribose. A general base-like role was also suggested, but in the absence of the identification of the nucleophile attacker atom, this role could not have been clarified. In the present structures, one of the side chain oxygens of Asp<sup>90</sup> (O2) provides close coordination to the catalytic water, and Asp<sup>90</sup>O<sub>H</sub> is also nearby (Fig. 4A).

Importantly, the present structural data also show that the proper orientation of the critical Asp<sup>90</sup> side chain is provided by an interaction between one of its other carboxyl oxygens (O61) and the main chain NH of the conserved Ala<sup>29</sup> from the neighboring subunit. This connection is strictly retained in all other dUTPase structures, as well, but no importance was ascribed to it. It also explains why the Asp<sup>90</sup>→Asn mutant is so much attenuated in catalytic efficiency and ineffective in coordinating the $W_{cat}$. In the mutant, the Ala<sup>29</sup> main chain NH retains its interaction with O61 of the Asn<sup>90</sup> residue, leaving only the nitrogen atom of the mutant (N82) for the role of close coordination to the catalytic water molecule. The nitrogen atom, however, cannot fulfill this role efficiently.

We also identified a second protein atom ligand of $W_{cat}$. The main chain carbonyl oxygen of Leu<sup>89</sup>, a conserved hydrophobic residue together with Asp<sup>90</sup> creates a well defined binding site for this water that is the second substrate of the enzyme.

At last, significance of Gln<sup>119</sup> (Motif IV) has also been discovered in this study. Its side chain oxygen (Oε1) coordinates a water molecule (W1) that provides H-bonding in the water network around $W_{cat}$. The Gln<sup>119</sup> side chain amino group participates in H-bonding to one of the oxygens on the reaction center α-phosphorous.

The Character of Phosphate Ester Hydrolysis—Available data for non-enzymatic phosphate diester hydrolysis reactions are in favor of a mechanism with significant associative character, since the metaphosphate transition state required for the dissociative mechanism is highly unstable when an additional bulky ligand is present (45, 46). Our present results for the dUTPase-catalyzed reaction are in agreement with this expectation. In our structure, the distance between the entering oxygen and the phosphorus reactive center is 3.6 Å with an experimental error of 0.16 Å based on the Luzzati plot (47), while a minimum of 4.9 Å is considered to be required for the dissociative mechanism (cf. Ref. 48). The present data show co-linearity of $W_{cat}$ with the scissile bond for in-line attack that can also be reconciled with a mechanism of significant associative character.

Acknowledgments—We gratefully thank Profs. Gábor Náray-Szabó and Richard I. Gumpert for fruitful discussions. We also thank Dr. Alexander Popov and Katja Schirzwitz for helpful suggestions in crystallization and x-ray diffraction, Imre Zagyva for excellent technical assistance, Rebecca Persson for the expression plasmid, and beamlines and operators at European Molecular Biology Laboratory/German Electron Synchrotron Facility (DESY) (Hamburg, Germany) and ID14-4 European Synchrotron Radiation Facility (Grenoble, France) for data collection opportunities.

REFERENCES

1. Friedberg, E., Walker, G., and Siede, W. (1995) DNA Mutagenesis and Repair, pp. 108–135, ASM Press, Washington, D. C.

2. Slhomai, J., and Kornberg, A. (1978) J. Biol. Chem. 253, 3305–3312

3. Pearl, L. H., and Savva, R. (1996) Nat. Struct. Biol. 3, 485–487

4. Ahern, G. W., and Browne, S. (1999) in Anticancer Development Guide: Antifolate Drugs in Cancer Therapy (Jackman, A. L., ed) pp. 409–421, Humana Press, Inc., Totowa, NJ

5. Canman, C. E., Radien, E. H., Parsels, I. A., Davis, M. A., Lawrence, T. S., and Maybaun, J. (1994) Cancer Res. 54, 2286–2288

6. Pugacheva, E. N., Ivanov, A. V., Krauczensko, I. E., Kopin, B. P., Levine, A. J., and Chumakov, P. M. (2002) Oncogene 21, 4595–4600

7. Benkovs, J. S., and Hammes-Schiffer, S. (2003) Science 301, 1196–1202

8. Knowles, J. R. (1991) Nature 350, 121–124

9. Schowen, R. L. (1978) in Transition States of Biochemical Processes (Gandour, R. D., and Schowen, R. L., eds) pp. 77–113, Plenum Press, New York

10. Dauter, Z., Persson, R., Rosengren, A. M., Nyman, P. O., Wilson, K. S., and Cedergren-Zeppezauer, E. (2001) Acta Crystallogr. D. Biol. Crystallogr. 57, 767–774

11. Larsen, G., Svensson, L. A., and Nyman, P. O. (1996) Nat. Struct. Biol. 3, 532–538

12. Larsson, G., Nyman, P. O., and Kvasman, J. O. (1996) J. Biol. Chem. 271, 10010–10016

13. Mustafi, D., Bekesi, A., Vertessy, B. G., and Makinen, M. W. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 5670–5675

14. Vagin, A., Barabas, O., Teplyakov, A., Bekesi, A., Dubrovay, Z., Pongracz, V., Zagyva, I., Imre, T., Szabo, P., and Vertessy, B. G. (2004) J. Biol. Chem. 279, 17932–17944

15. Persson, T., Larsson, G., and Nyman, P. O. (1996) Bioorg. Med. Chem. 4, 553–556

16. Persson, R., Nord, J., Roth, J., and Nyman, P. O. (2002) Prep. Biochem. Biotechnol. 32, 157–172

17. Vertessy, B. G. (1997) Proteins 28, 568–579

18. Prasad, G. S., Stura, E. A., Elder, J. H., and Stout, C. D. (2000) Acta Crystallogr. D. Biol. Crystallogr. 56, 1100–1109

19. Persson, R., Nord, J., Roth, J., and Nyman, P. O. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 38812–38817

20. Vagin, A., and Teplyakov, A. (1997) J. Appl. Crystallogr. 30, 1022–1025

21. Taylor, J. S. (1981) J. Biol. Chem. 256, 1510–1511

22. Ashburner, M., and D 설치용, M. L., and Brundiers, R., eds) pp. 77–113, Plenum Press, New York

23. Dauter, Z., Persson, R., Rosengren, A. M., Nyman, P. O., Wilson, K. S., and Cedergren-Zeppezauer, E. (2001) Acta Crystallogr. D. Biol. Crystallogr. 57, 767–774