J|A|C|S
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

Structural Characterization of Arginine Fingers: Identification of an Arginine Finger for the Pyrophosphatase dUTPases

Gergely N. Nagy,*,[†,‡] Reynier Suardíaz,[§] Anna Lopata,[‡,∇] Olivér Ozohanics,[‖] Károly Vékey,[⊥,○] Bernard R. Brooks,[#] Ibolya Leveles,[†,‡] Judit Tóth,[‡] Beata G. Vértessy,*,[†,‡] and Edina Rosta*,[§]

[†]Department of Biotechnology and Food Sciences, Budapest University of Technology and Economics, Budapest 1111, Hungary

[‡]Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest 1117, Hungary

[§]Department of Chemistry, King's College London, London SE1 1DB, United Kingdom

[‖]MS Proteomics Research Group, Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest 1117, Hungary

[⊥]Core Technologies Centre, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest 1117, Hungary

[#]Laboratory of Computational Biology, National Heart, Lung and Blood Institute, National Institutes of Health, Rockville, Maryland 10892-9314, United States

S *Supporting Information*

ABSTRACT: Arginine finger is a highly conserved and essential residue in many GTPase and AAA+ ATPase enzymes that completes the active site from a distinct protomer, forming contacts with the γ-phosphate of the nucleotide. To date, no pyrophosphatase has been identified that employs an arginine finger fulfilling all of the above properties; all essential arginine fingers are used to catalyze the cleavage of the γ-phosphate. Here, we identify and unveil the role of a conserved arginine residue in trimeric dUTPases that meets all the criteria established for arginine fingers. We found that the conserved arginine adjacent to the P-loop-like motif enables structural organization of the active site for efficient catalysis via its nucleotide coordination, while its direct electrostatic role in transition state stabilization is secondary. An exhaustive structure-based comparison of analogous, conserved arginines from nucleotide hydrolases and transferases revealed a consensus amino acid location and orientation for contacting the γ-phosphate of the substrate nucleotide. Despite the structurally equivalent position, functional differences between arginine fingers of dUTPases and NTPases are explained on the basis of the unique chemistry performed by the pyrophosphatase dUTPases.
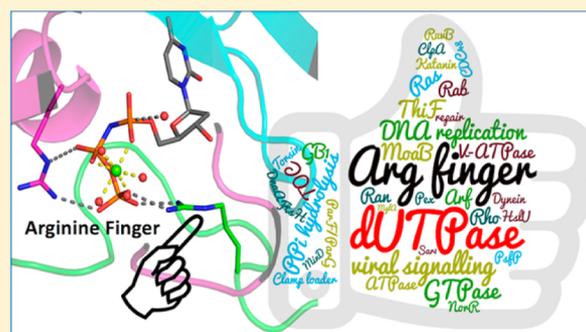
INTRODUCTION

Protein–protein interactions play crucial roles in ensuring optimal catalytic activity. Notably, interprotomer interactions often modulate the nucleotide triphosphate (NTP) hydrolysis and transfer reaction rates in NTP processing enzymes, e.g., AAA+ enzymes[1] (ATPases associated with various cellular activities), GTPases,[2] polymerases,[3] kinases,[4,5] and dUTPases[6] (DUT) assembled as homo- or heteropolymers. In many of these enzymes, an essential conserved residue, the arginine (Arg) finger, has been identified as a key entity in catalytic rate acceleration. The Arg finger constitutes a catalytically active complex with a neighboring enzyme monomer through interactions with the γ-phosphate of the nucleotide substrate.[7] The Arg finger can originate either from a neighboring enzyme protomer (e.g., AAA+ enzymes,[8] large GTP binding proteins[9]) or from an effector protein (e.g., small GTPases[7]). This residue was proposed to possess a catalytic role via nucleotide polarization and/or transition state (TS) stabilization in GTPases,[7,10] ATPase motors,[11,12] and helicases,[13] and in signal transduction ATPases.[14] Experimental data established that

even the conservative replacement of the Arg finger[7,15] diminished the catalytic rates while maintaining local structural integrity including P-loop coordination.

All proteins with catalytically essential Arg fingers reported in the literature to date catalyze the hydrolysis at the γ-phosphate group of the NTP. However, highly conserved arginines also assist the mechanism of NTP pyrophosphatases or transferases. Among them, dUTPases catalyze the pyrophosphorolysis of dUTP, a noncanonical DNA nucleotide.[16] With this catalytic activity, dUTPases participate in preventive DNA repair and were recently shown to also exhibit regulatory functions related to, e.g., viral gene expression.[17,18] Trimeric dUTPases host three active sites formed with the essential contribution of each protomer[19] (Figure 1a). The nucleotide interactions at each active site are provided by five conserved motifs. The C-terminal arm of trimeric dUTPases encompasses the dynamic fifth conserved motif,[20] which contains a conserved
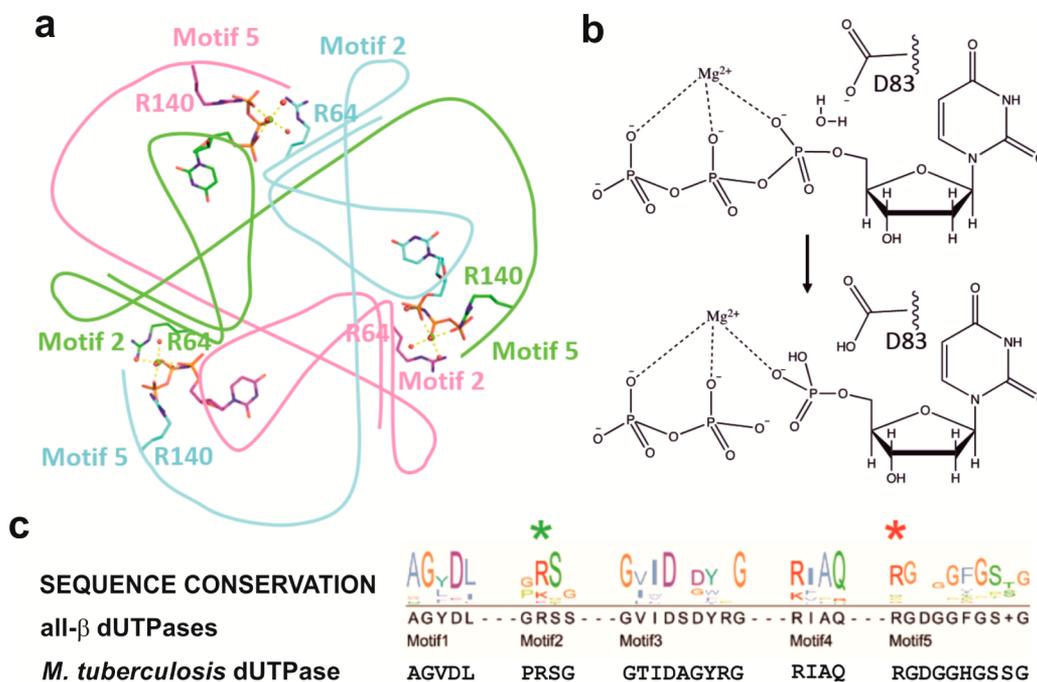
Figure 1. Structural organization and catalytic mechanism of trimeric dUTPases. (a) Schematic representation of the trimer assembly and active site organization. Conserved arginine residues (R64 and R140) that coordinate the nucleotide triphosphate together with the catalytic Mg^{2+} ions and coordinating water are highlighted (licorice and spheres). (b) Reaction scheme for the rate limiting nucleotide hydrolysis reaction step. Nucleophilic attack on the Pα atom is supported by the conserved proton-abstracting residue D83. (c) Sequence conservation of motifs 1–5 in all-β folded (trimeric or monomeric) dUTPases. The conserved arginines of 2nd and 5th motifs are highlighted with green and red asterisks, respectively. The mtDUT sequence is also shown for comparison.

Arg directly preceding a glycine-rich P-loop-like motif (Figure 1). This Arg residue forms multiple hydrogen-bonding contacts with the γ-phosphate of the substrate and contributes to the functionally critical dUTP/dUDP substrate discrimination of dUTPases.[20,21] Despite the wealth of biochemical, structural, and computational studies concerning dUTPases,[20,22−27] structural and computational evidence assessing the role of this conserved Arg was missing.

Here, we determined crystallographic structures of *Mycobacterium tuberculosis* dUTPase (mtDUT), to establish the role of R140 as an arginine finger, performing kinetic and ligand binding measurements as well as molecular dynamics (MD) and QM/MM calculations. Our results together reveal the essential role of R140 in catalysis by primarily promoting active site organization. We further provide an exhaustive structure-based comparison of analogous conserved arginine residues present in a large number of nucleotide hydrolase and transferase enzyme classes, identifying a conserved spatial position for these positively charged Arg residues in one metal-ion catalytic[28] nucleotide hydrolase and transferase enzymes.

■ METHODS

Chemicals. DpnI restriction enzyme as well as Phusion polymerase were obtained from New England Biolabs (Ipswich, MA). IPTG was obtained from Fisher Scientific GmbH, Germany. Ni-NTA was from Qiagen. Materials for electrophoresis were from Bio-Rad. Phenol red was obtained from Merck (Kenilworth, NJ); dUTP, α,β-imido dUTP (dUPNPP), and dUDP were from Jena Biosciences. All other chemicals were of analytical grade and were purchased from Sigma-Aldrich (St. Louis, MO).

Mutagenesis and Protein Expression. Site-directed mutagenesis was performed by the Quikchange method (Agilent) and was verified by sequencing (MWG Eurofins GmbH, Martinsried, Germany). We interrogated the effect of a lysine surrogate to the R140 residue by performing the arginine/lysine exchange in mtDUT^{H145W} construct which displays similar kinetic parameters to the wild type mtDUT.[29] The following primers were used to create mtDUT^{R140K,H145W} from the mtDUT^{H145W} construct: 5′−3′ GGCCTCGAC-ATCC<u>AAG</u>GGCGACGGTGGCTGGGG and 3′−5′ CCCCAGCCACCGTCGCC<u>CTT</u>GGATGTCGAGGCC, respectively. For mtDUT^{G143STOP}, the primers 5′−3′ CCCGCGGCGACT-GAGGCCACGGTTCCTC and 3′−5′ GAGGAACCG-TGGCC<u>TCA</u>GTCGCCGCGGG enabled the introduction of a stop codon in place of G143 residue of mtDUT^{H145W}, thus disrupting the first β-turn of the P-loop-like motif. Proteins were expressed and purified as described previously.[29] Briefly, recombinant dUTPase with an N-terminal 6xHis-tag was expressed in *E. coli* strain BL21(DE3)-pLysS. The final supernatant after cell extraction was loaded onto a Ni-NTA column and purified according to the Novagen protocol. Proteins were dialyzed against a buffer containing 20 mM HEPES/NaOH pH 7.5, 100 mM NaCl, 2 mM MgCl$_2$, and 1 mM DTT. Protein concentration was measured by UV absorbance using extinction coefficients ε_{280} = 8480 M^{-1} cm^{-1} for mtDUT^{R140K,H145W} and ε_{280} = 2980 M^{-1} cm^{-1} for mtDUT^{G143STOP}.

ESI-MS. In the mass spectrometric study of protein complexes, a commercial Waters QTOF Premier instrument (Waters, Milford, MA) equipped with an electrospray ionization source (Waters, Milford, MA) was used in the positive ion mode. Mass spectra were obtained under native conditions; the ions were generated from aqueous 10 mM NH$_4$HCO$_3$ pH 7.5 buffer solution containing *M. tuberculosis* dUTPase protein constructs at 0.4 μM monomer concentration. These conditions allowed transfer of the native protein complex present in the solution into the gas phase. The capillary voltage was 3600 V; the sampling cone voltage was 125 V, and the temperature of the source was kept at 80 °C. Collision cell pressure was 3.38 × 10^{-3} mbar, and ion guide gas flow was 35.00 mL/min. Mass spectra were recorded using the software MassLynx 4.1 (Waters, Milford, MA) in the mass range 1400−5000 m/z.

Enzyme Activity Assay. dUTPase activity was followed by a continuous spectrophotometric phenol red assay as described in ref 30.

Table 1. Kinetic and Ligand Binding Parameters of the dUTPase Variants

enzyme	mtDUT^{WT}	mtDUT$^{H145W a}$	mtDUT$^{T138STOP a}$	mtDUT^{R140K,H145W}	mtDUT^{G143STOP}
k_{cat} (s^{-1})	3.0 ± 0.3	0.7 ± 0.04	0.0035 ± 0.0001	0.0045 ± 0.0003	0.00089 ± 0.00004
K_M (μM)	1.4 ± 0.4	1.3 ± 0.4	6.7 ± 0.4	1.1 ± 0.3	8.2 ± 2.5^b
$K_{ddUPNPP}{}^b$ (μM)	n.d.	0.9 ± 0.5^a	3.9 ± 1.3	5.7 ± 1.5	5.2 ± 1.0

[a]From ref 24. [b]K_d measurements were performed by near-UV CD titration, as described in the Methods section; for titration curves, see also Figure S2.

Briefly, the assay that measures proton release upon dUTP hydrolysis was performed in 10 mm path length cuvettes thermostated to 20 °C in a Specord 200 spectrophotometer with a reaction volume of 1000 μL. Due to low enzyme activity of the investigated enzyme point mutants, mtDUT^{R140K,H145W} and mtDUT^{G143STOP} enzyme concentrations were set to 1 and 5 μM, respectively, in 1 mM HEPES pH 7.5 buffer containing 100 mM KCl, 40 μM phenol red, and 5 mM MgCl$_2$. The reaction was started by a final quick addition of 5−80 μM dUTP to the reaction mixture; then, absorbance decrease was recorded for 1−5 min. The measured absorbance decline was corrected for the change of the absorbance for the same time interval in absence of dUTP substrate. The average of three measurements was calculated at each data point. Kinetic data were fitted with Michaelis−Menten equation.

Circular Dichroism Intensity Titrations. Near-UV CD spectra were recorded at 20 °C on a JASCO 720 spectropolarimeter using a 10 mm path length quartz cuvette. Protein (50 μM) was titrated with stepwise addition of α,β-imido-dUTP (dUPNPP) in a buffer containing 10 mM potassium phosphate pH 7.5 and 1 mM MgCl$_2$, as described in ref 20. A spectrum in the range λ = 250−290 nm was recorded at each nucleotide concentration. Differential curves were obtained by subtracting the CD spectra of nucleotide−buffer titrations from that of the corresponding nucleotide−protein titration point. Differential ellipticity at λ_{max} = 269 nm was plotted against the nucleotide concentration by using Origin software (OriginLab Corp., Northampton, MA). Dissociation constants were determined by fitting the data with quadratic eq 1:

$$y = s + \frac{A[(c + x + K) - \sqrt{(c + x + K)^2 - 4cx}]}{2c} \qquad (1)$$

Here $s = y$ at $x = 0$; A is the amplitude of the circular dichroism intensity change. c is the concentration of the protein, and K is the dissociation constant.

Crystallization. mtDUT^{R140K,H145W} and mtDUT^{G143STOP} proteins were crystallized using the hanging drop method and applying the same crystallization condition that was used for the wild type enzyme.[29] Briefly, 4.5 mg/mL dUTPase together with 6.7 mM dUPNPP, 20 mM MgCl$_2$, and 0.1% sodium azide were mixed with the reservoir solution containing 0.1 M TRIS/HCl pH 7.5, 1.5 M ammonium sulfate, and 12.5% glycerol in a 1 μL:1 μL ratio.

Data Collection and Structure Determination. Complete crystallographic data sets of the acquired crystals were recorded at a Supernova sealed tube home source diffractometer (Agilent) for mtDUT^{R140K,H145W}, as well as at the ID23-2 beamline of ESRF (Grenoble), for mtDUT^{G143STOP}. Data were processed using XDS and XScale[31] programs. An initial electron density map was obtained for both mutant enzyme:dUPNPP:Mg^{2+} complexes by Fourier synthesis calculating with the phases of the mtDUT^{H145W} structure (PDB: 3HZA) as the model. Model building and refinement was carried out using the Coot[32] program and REFMAC[33] from the CCP4 suite.[34] To model anisotropic displacements of the protein complexes, TLS parametrization was implemented within the program REFMAC. Later, refinement of high resolution mtDUT^{G143STOP} structure was pursued using anisotropic B factors. At the final cycles of this refinement, B factor restraints were reduced to 0.7. Final models contain the bound substrate analogue dUPNPP:Mg^{2+} complex. For the sake of clarity, residues are numbered according to the physiological *M. tuberculosis* dUTPase sequence, disregarding the recombinant tag in the atomic coordinate files as well as in figures and

throughout the text. Coordinates and structure factor data of mtDUT^{R140K,H145W} and mtDUT^{G143STOP} were deposited to the Protein Data Bank with accession codes 5EDD and 5ECT, respectively (cf. Table 2).

Molecular Dynamics Simulations. The coordinates of the wild type (WT) mycobacterial dUTPase enzyme in complex with the isosteric substrate analogue, α,β-imido-dUTP, and Mg^{2+} were taken from the PDB 2PY4 structure of 1.49 Å resolution. To convert the crystallized substrate analogue to dUTP, the imido NH moiety was replaced by O. The R140K and G143STOP mutants were generated using the Mutagenesis utility of Pymol[35] starting from the wild type 2PY4 structure. The simulations were set-up using standard simulation protocols with CHARMM-GUI.[36,37] Hydrogen coordinates were generated using standard protonation states for all ionizable residues using CHARMM.[38] The system was solvated with a pre-equilibrated TIP3P cubic water box of 80 Å^3. Water molecules were randomly replaced by ions to ensure the neutralization of the system, and an additional KCl salt concentration corresponding to 0.15 M. We run production trajectories of 300 ns after 10 ns of equilibration using NAMD.[39] Temperature and pressure were held constant at 300 K and 1 atm, respectively. We used the CHARMM36[40] force field with periodic boundary conditions and the particle mesh Ewald method[41] for the long-range electrostatics in combination with a 12 Å cutoff for the evaluation of the nonbonded interactions, and the same protocol was followed for the corresponding MD simulations.

QM/MM Calculations. Geometries of the reactant and transition states of the phosphate cleavage reaction in WT dUTPase were taken from our previous work.[24] The system was trimmed to a sphere of 21 Å radius centered at the Pα atom of the dUTP. Atoms further than 15 Å from the Pα atom of dUTP were kept frozen during all QM/MM calculations. No cutoffs were introduced for the nonbonding MM and QM/MM interactions. We have used Q-Chem v4.1[42,43] and Gaussian09 D01[44] program packages to perform the QM calculations using the B3LYP functional in combination with 6-31+G(d,p) basis sets. The quantum mechanical system was coupled with the CHARMM program[38] for the MM region. A full electrostatic embedding scheme[45] has been adopted in all the calculations, and hydrogen link atoms have been used to treat the QM/MM boundaries. The energy minimized reaction pathway was obtained using a QM region consisting of the dUTP molecule; the backbone of T81 and I82; the side chains of D83, S65, H145, S147, and S148; the amide group of Q113; the guanidine group of R140; the Mg^{2+} ion with its full coordination shell; and nearby crystallographic waters to account for 132 atoms. We investigated the reaction barrier using different QM/MM model systems. To separate the geometrical factors from the electrostatic and charge transfer interactions, we used the fully geometry optimized WT reaction profile,[24] and recalculated the barrier height using different models for the R64 and R140 side chains. We included them either (i) in the QM region, corresponding to the minimized profile, accounting fully for charge transfer and electrostatic effects; (ii) in the MM region, allowing only electrostatic effects; or (iii) with its atomic charges set to zero (neutralized) during the evaluations of the QM/MM wave function and energies, only preserving its geometrical effects.

■ RESULTS

Integrity of Both R140 and the P-Loop-like Element Is Essential for Catalytic Efficiency. We engineered two mutant proteins (mtDUT^{R140K,H145W} and mtDUT^{G143STOP}) to

Table 2. Crystallographic Data Collection and Refinement Statistics

	mtDUT^{R140K}	mtDUT^{G143STOP}
PDB	5EDD	5ECT
	Data Collection	
space group	$P6_3$	$P6_3$
cell dimensions a, c (Å)	55.05, 83.83	54.78, 84.08
resolution (Å)	41.9−1.97 (2.02−1.97)[a]	26.0−1.30 (1.33−1.30)[a]
wavelength (Å)	1.5406	0.8729
R_{merge}	0.130 (0.497)	0.028 (0.561)
$\langle I/\sigma(I) \rangle$	6.1 (1.6)	21.23 (1.95)
completeness (%)	99.6 (99.6)	99.1 (99.3)
redundancy	3.5 (2.5)	3.3 (3.3)
	Refinement Statistics	
resolution (Å)	19.19−1.97	
no. reflns	35829	34820
R_{work}/R_{free}	0.160/0.211	0.123/0.157
no. atoms		
protein	1026	1070
ligand/ion	37	39
water	159	127
B factors		
protein	21.0	22.3
ligand/ion	20.2	19.1
water	32.0	31.6
rms deviations		
bond length (Å)	0.019	0.019
bond angle (deg)	1.95	1.76

[a]Values in parentheses are for the highest resolution shell.
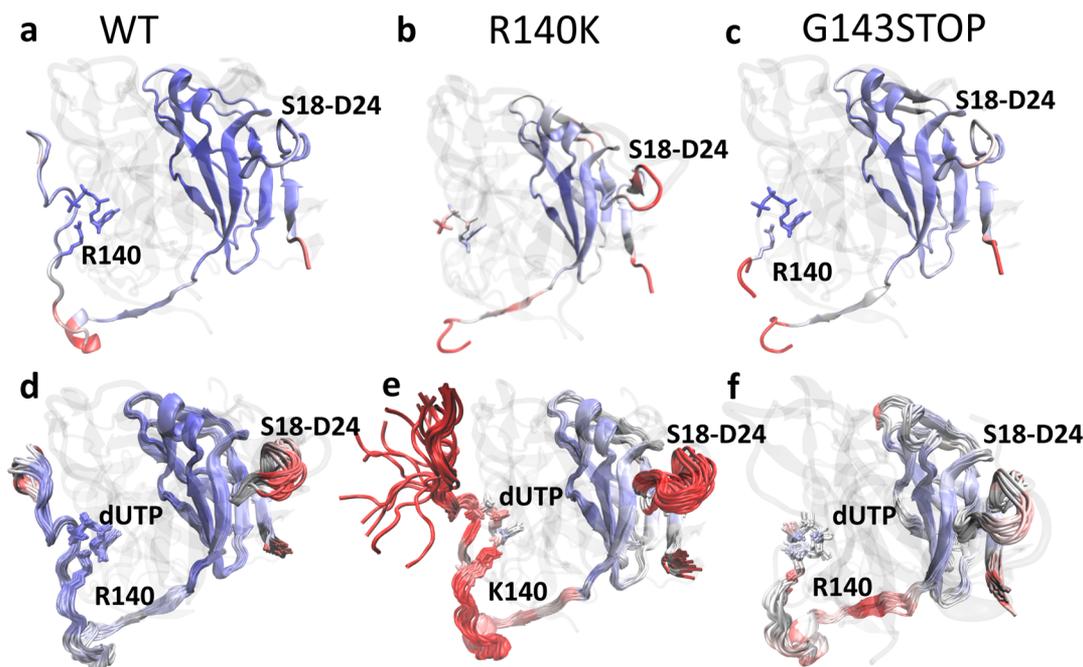
Figure 2. Flexibility and disorder of mtDUT and its mutants affecting the Arg finger and the P-loop-like elements from crystallographic structures (top) and MD simulations (bottom). Cartoon representation is shown of the (a) wild type mtDUT (PDB: 2PY4), (b) mtDUT^{R140K,H145W}, and (c) mtDUT^{G143STOP} structures. The local thermal motions defined by B factors in crystal structures (upper row, a−c) are compared with the observed flexibility in MD simulations (lower row, d−f). The superposition of MD simulation snapshots collected at uniform intervals is shown. The disordered C-terminal region and the loop S18−D24 are highlighted; the dUTP and the R/K140 are shown as licorice. Color code: blue, more rigid; red, more flexible. Only one dUTPase protomer is highlighted with the other two protomers shown as transparent.

eludicate the role of key constituents of the C-terminal arm including the R140 in substrate coordination and catalysis. Notably, the conserved arginine is replaced by a lysine in the first construct, whereas the C-terminal truncation of the second construct at the 143th residue position (mtDUT^{G143STOP}) disrupts only the P-loop-like element, while the R140 is

retained in the sequence. The H145W mutation introduces an intrinsic Trp fluorescence sensor that we have extensively used for the characterization of the enzymatic mechanism of various dUTPases.[29,46,47] The mtDUT[H145W] enzyme was reported to behave as wild type.[29] Both mtDUT[R140K,H145W] and mtDUT[G143STOP] form homotrimers, as confirmed by ESI-MS (Figure S1, Table S1). Steady-state kinetic measurements revealed severely deteriorated dUTPase activities for these constructs with corresponding k_{cat} values comparable to that of mtDUT[T138STOP], devoid of the P-loop-like element as well as of R140[24] (Table 1, Figure S2). Near-UV circular dichroism equilibrium ligand binding measurements showed that both mtDUT[R140K,H145W] and mtDUT[G143STOP] have a moderate 5-fold decreased affinity toward the slowly hydrolyzable substrate analogue α,β-imido dUTP (dUPNPP) (Table 1, Figure S2).

High Resolution mtDUT Crystal Structures Provide Evidence That R140 Is Critical for Proper Active Site Organization. We determined the mtDUT[R140K,H145W] and mtDUT[G143STOP] crystal structures to 1.97 and 1.30 Å resolution, respectively (Table 2). These structures display similar overall fold as the mtDUT[WT] (PDB: 2PY4)[29] with all atom rms deviations of 0.172 and 0.174 Å, respectively. In both mutant structures, the dUPNPP:Mg^{2+} complex is clearly visible in the enzyme active site in a conformation similar to that in the wild type enzyme (Figure S3). In addition, an electron density is present in both mutant structures at the approximate position of the catalytic water molecule. Strikingly, the complete C-terminal arm, encompassing the K140 and the P-loop-like element, becomes disordered in the R140K mutant (Figure 2, Figure S3). These alterations render the highly conserved R64 to be the sole residue that still directly interacts with the dNTP γ-phosphate. The structure of the mtDUT[G143STOP] also lacks the P-loop-like element, simply due to the engineered truncation. Nevertheless, the electron density corresponding to R140 together with vicinal residues (138−141) is resolved, despite the C-terminal arm segment 134−137 not being resolved in this structure. Here, the bidentate interaction of R140 with the nucleotide γ-phosphate persists. However, R64, the other interaction partner of the nucleotide γ-phosphate, is shifted by 0.9 Å away from its wild type location, contributing to a twisted conformation of the triphosphate chain (Figure S3). In both structures, the abolished conserved interactions of the nucleotide with the P-loop-like element are substituted by ordered water molecules (Figure S3). In accordance with the loss of important interactions, the triphosphate groups of the nucleotides are characterized with an increased flexibility as deduced from B factor analysis (Figure S4).

Alteration of R140 Affects Ligand Traffic. Crystallographic B factor analysis of mtDUT[R140K,H145W] revealed that the loop (S18−D24) between the first and the second β-sheets directly adjacent to the first conserved motif appears with increased flexibility compared to mtDUT[WT] (Figure 2). A comparison with other deposited mtDUT structures revealed that increased flexibility of the loop (S18−D24) is induced by exchange or deletion of R140, which engages in an interprotomer interaction with D24 (Table S2). This contact together with the H21−S139 interaction is of key importance to the ordering of the C-terminal arm to the active site as the preceding part of the arm (residues 130−138) is not involved in polar interactions with other protein segments. This anchoring of R140 further facilitates interactions of the P-loop-like element (141−148) with the triphosphate chain of the nucleotide. Molecular dynamics simulations revealed that

the dynamic movements of this loop facilitate the entry and departure of the nucleotide to the active site.[48]

MD Simulations and QM/MM Calculations Reveal the Role of Arginine Residues in the Catalytic Mechanism of dUTPase. MD simulations were performed for the mtDUT[R140K], mtDUT[G143STOP] as well as for the mtDUT[WT] trimer structures with dUTP substrate present at the active sites. Each simulation was initiated using the wild type structure as template, and was run for 300 ns. In the case of the mtDUT[R140K] mutant, the C-terminal arm became flexible in two of the three active sites of the mtDUT[R140K] trimer within the first 10 ns, whereas the third active site became disordered only after about 150 ns, reflected by the distance between K140 and the γ-phosphate (Figure 2e, Figure S5). The disorder at the C-terminal region is accompanied by the increased disorder of the S18−D24 loop of the interacting protomer, indicating extensive local structural perturbations. In contrast, MD simulations of mtDUT[G143STOP] indicated low flexibility for the available part of the C-terminal arm together with unaltered γ-phosphate coordination by R140 (Figure 2f, Figure S5). Nevertheless, the interactions with the other highly conserved arginine, R64, and the nucleotide γ-phosphate became looser, in accordance with the structural results (Figures S3 and S5). To probe the significance of R64 and R140 coordination on the catalytic mechanism of dUTPase, we calculated the reaction barrier heights of the pyrophosphate cleavage reaction at the QM/MM level using several model systems. We found that the electrostatic influence of R140 in the catalytic step is only minor (Figure 3 and Figure S6). The change in the activation energy was within 2 kcal/mol, regardless of whether the R140 side chain was present in the QM region or in the MM region, or if it was neutralized. In contrast, the absence of R64 from the MM region rendered the enzyme virtually inactive with a reaction barrier increase by more than 10 kcal/mol (Figure 3 and Figure S6). Including R64 in the QM region did not change the reaction barrier considerably. Therefore, we suggest that the key effect of this latter residue on the catalysis is mainly electrostatic, whereas R140 does not have a major electrostatic stabilization effect for the transition state.

Structural Comparison of Arg Finger−NTP Interactions for All Enzyme Classes. To assess the general structural arrangement of Arg−NTP interactions, we analyzed the positions of Arg residues from a large number of enzyme classes (Figure 4) that form hydrogen bonds with a nucleoside triphosphate substrate from high resolution crystallographic structures (Figure S7). We selected enzyme structures with a single bound Mg^{2+} cofactor in the active site, which is a key common feature among dUTPases, GTPases, or AAA+ enzymes, and for all of the known Arg finger examples (Figure 5), to the best of our knowledge. The structural alignment identifies a large Arg cluster near the γ-phosphate, with ATPases and GTPases as primary examples with an Arg finger residue (Figure 4, Figure S7). A second, more scattered smaller cluster corresponds to various kinases and the R64 of dUTPase (Figure 4a). These results suggest that an arginine residue positioned to contact the γ-phosphate at the specific orientation as shown by the main cluster is generally advantageous for the γ-phosphate catalytic cleavage reactions. The fifth motif R140 in mtDUT protomers also occupies this equivalent position. Nevertheless, an important functional difference arises from the fact that dUTPases catalyze a chemically different catalytic reaction as compared with GTPases and ATPases (Figure 4). For trimeric and monomeric dUTPases, the nucleophilic attack
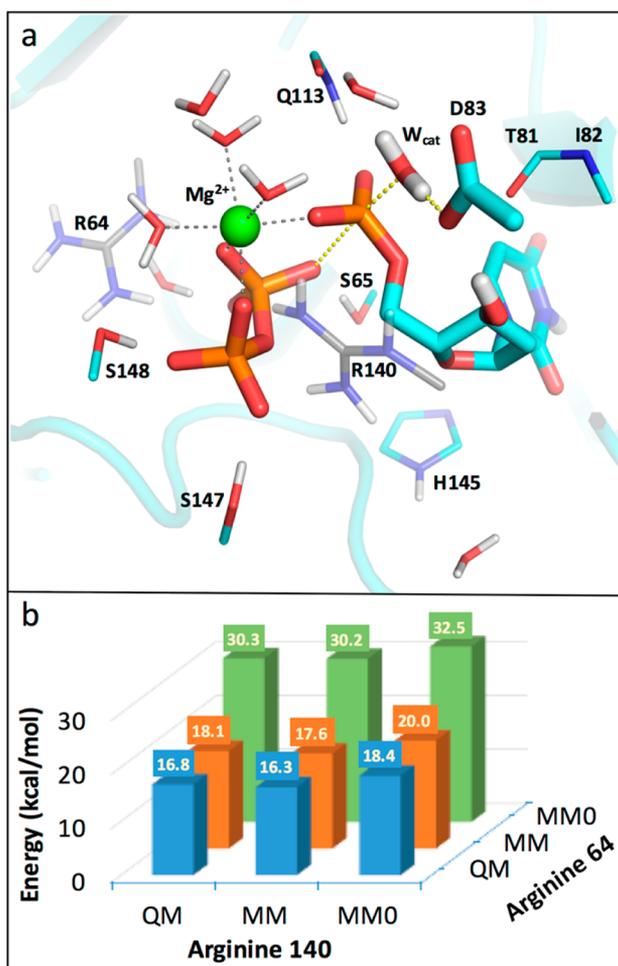
Figure 3. Role of arginine residues in mtDUT catalysis as assessed by QM/MM calculations. (a) Transition state configuration of the *M. tuberculosis* dUTPase active site. Residues always included in the QM region (licorice, sphere for the Mg^{2+}) and the R64 and R140 moieties (transparent thin sticks) are shown. The reaction coordinate was defined as a linear combination of distances involving the breaking and forming of bonds (yellow dashed lines): $r(O−P\alpha) − r(Ow−P\alpha) + 0.5r(Ow−Hw) − 0.5r(Hw−O_{C=O})$. (b) QM/MM energy barriers for the catalytic step when the side chains of R64 and/or R140 are (i) included in the QM region (QM), (ii) included in the MM region (MM), or (iii) neutralized in the MM region (MM0) during QM/MM calculations.

occurs at the α-phosphate; therefore, their cognate Arg corresponding to R140 is not involved in the direct electronic stabilization of the transition state. In contrast, the γ-phosphate serves as the reaction center upon the nucleophilic attack in GTPases and ATPases, in which the Arg fingers can indeed stabilize the transition state.

■ DISCUSSION

Functional Roles of Conserved Arginines at the Active Site of dUTPases. Three conserved arginine residues are present in the active site of dUTPases: R64 (second motif), R110 (fourth motif), and R140 (fifth motif). Both R64 and R140 directly contact the triphosphate chain (Figure 1c). Arg140 coordinates the nucleotide γ-phosphate, analogously to the Arg fingers in AAA+ enzymes. The fourth motif R110 does not directly coordinate the substrate nucleotide, and the corresponding R111A mutant of *Saccharomyces cerevisiae*

dUTPase showed only 30−50% decrease in activity.[27] R64 contacts both the β- and γ-phosphates (Figure 4b), and at the same time completes the active site at the interface between two protomers. R64 is part of a different protomer than the one coordinating the nucleoside group. In contrast to R140, our QM/MM results predicted that R64 electrostatically stabilizes the TS of the pyrophosphorolysis reaction. This is also in accordance with a previous mutational study in *Saccharomyces cerevisiae* that reported very low activity for the corresponding Ala mutant R68A.[27] Furthermore, a positively charged residue (R/K) occupies this position almost invariably, and the highly conserved R64 is substituted by Lys in only about 3% of fully functional trimeric dUTPases (e.g., ref 22) (Figure 1c and Table S2).

R140 is a highly conserved residue among trimeric and monomeric dUTPases (Figure 1b). Its naturally occurring homologues are very rare (less than 0.3%, see Table S3), and do not favor the conservative Lys exchange. This is in good agreement with our results that conservative R140K mutation renders the enzyme inactive. Accordingly, Arg/Lys, Arg/Ala point mutants of dUTPases were also seriously compromised in their catalytic efficiency on the basis of prior studies including herpes viral, lentiviral, retroviral, yeast, and human sources.[20,23,25−27]

R140 displays an exceptional combination of sequential and steric positioning in dUTPases that enables its major effects on the P-loop-like motif coordination. R140 directly precedes the P-loop motif, atypical for NTPases with the conventional GxxxxGK(T/S) P-loop or Walker A type sequence motif.[50] Our results demonstrate that deletion of the P-loop-like element renders the *M. tuberculosis* dUTPase inactive. On the basis of our combined *in vitro*, computational, and structural data of mtDUT[R140K,H145W] and mtDUT[G143STOP], we propose different mechanisms for the activity loss of the two mutant enzyme constructs. In the presence of Arg140 and in the absence of the P-loop-like motif, the conserved Arg residue provides similar contacts to the nucleotide as in the wild type enzyme. Therefore, low enzymatic activity shows that R140 in itself cannot provide sufficient catalytic rate enhancement, and the loss of catalytically competent P-loop coordination is likely the main reason for the lack of activity. Accordingly, several earlier studies demonstrated that mutant constructs devoid of the P-loop region[23,51] or those lacking the complete C-terminal arm including the conserved Arg[20,22−24] are nearly inactive.

In contrast, when R140 is exchanged to a Lys, the deteriorated enzyme activity together with structure-based and MD observations reflects that Arg/Lys substitution is not functional at this position. Our results indicate that R140 residue is therefore an essential factor for the structural integrity of the active site to enable efficient catalysis: (i) via its direct nucleotide coordination or (ii) via interprotomer contacts. These interactions enable optimal positioning of hallmark residues from the P-loop-like element,[20,24,52] including S147 and S148 that have a decisive role in ensuring catalytic efficiency. This is also in agreement with QM/MM calculations demonstrating that the absence of the complete C-terminal arm[24] but not of R140 in itself corresponds to a marked larger increase of the transition state energy barrier, suggesting a potent yet indirect role of R140 in catalytic rate enhancement. It is worth noting that in our QM/MM calculations we used the optimized geometries along the profile obtained for the wild type reaction, and only single point calculations were performed for our model systems to distinguish between the
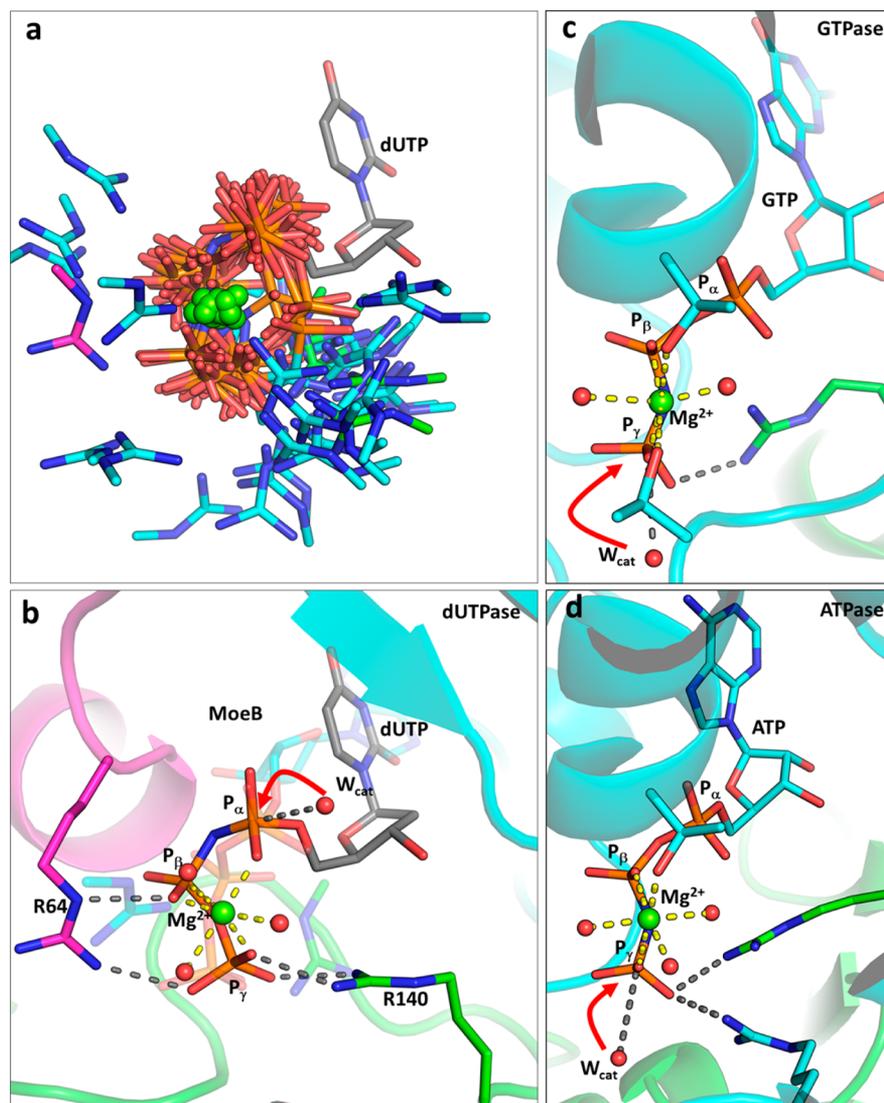
Figure 4. Arg fingers from various NTP hydrolase/transferase enzymes. (a) Arg residues occupy clusters in specific positions around the γ-phosphate in NTP cleaving enzymes. The nucleoside group (gray) is shown for dUTPase. The triphosphate group and its coordinating Arg residues are shown as sticks (blue for the same protomer as the one that coordinates the substrate, green or pink for additional different protomers). Mg^{2+} ions are shown as green spheres. (b) Analogous positions of the Arg fingers around the nucleotide for dUTPase (PDB: 2PY4,[29] note the presence of α,β-imido dUTP substrate analogue instead of dUTP) and for the MoeB protein (transparent, 1JWA[49]). (c, d) Analogous positions of the Arg fingers from a GTPase (PDB: 2QTV, c) and an ATPase (PDB: 2JIZ, d). Note that the nucleophilic attack (red arrows) of the respective nucleophile water molecule (W_{cat}) takes place on different phosphate moieties for the pyrophosphatase dUTPases and MoeB vs ATPases and GTPases.

electrostatic and structural contributions to catalysis. More detailed analysis could also be carried out to evaluate the electrostatic preorganization toward the transition state, as done, e.g., by Florian and Warshel.[53] To compare with our biochemical data, calculations of accurate free energy barriers are needed, taking into account geometrical relaxation that requires more extensive free energy calculations.

Identification of an Arg Finger in dUTPases. On the basis of the following properties, we propose that R140 fulfills the requirements of being an Arg finger: (i) It is a highly conserved residue. (ii) Its mutation (even conservative) seriously compromises the catalytic activity. (iii) It emerges from a separate protein subunit that forms the substrate binding site and establishes interactions with the γ-phosphate of the substrate NTP. The conserved arginine residue of dUTPases thus displays functional and structural similarities to the Arg finger motifs of AAA+ enzymes, GTPase activating proteins

(GAPs) for small G proteins, and other nucleoside triphosphate hydrolyzing enzymes that form interactions with the γ-phosphate of the substrate. For a more in-depth structural analysis, we provided a thorough structure-based comparison of arginines and, within, Arg fingers employed by NTP hydrolase enzymes (Figure 4). Importantly, while the P-loop-like element was identified in other pyrophosphatase enzymes as well,[54−56] they do not employ an Arg finger. Notably, conserved Arg residues are also present in two metal-ion catalyst enzymes[57] (e.g., dimeric dUTPases, polymerases); however, these residues originated from the same protein subunit as the protomer forming the active site. Therefore, these are not considered Arg fingers. Moreover, the positioning of the arginines in these enzymes does not seem to follow an obvious cluster around the triphosphate chain (Figure S8).

Interestingly, R64 also fulfils the requirements for an arginine finger, thus presenting dUTPases with two arginine fingers.
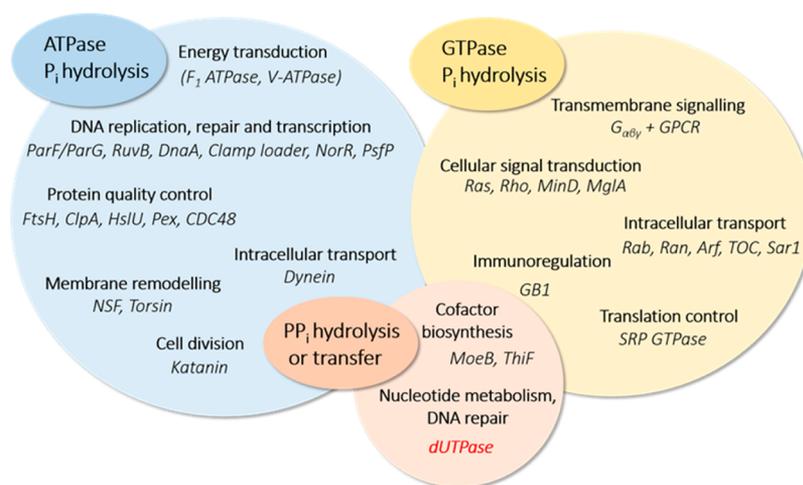
Figure 5. Biological role and biochemical function of enzymes employing arginine fingers. The diverse physiological functions fulfilled by ATPases, GTPases, and PP$_i$ hydrolysis or transfer enzymes are displayed together with enzyme representatives (for additional structural information, see Figure S7).

Dual essential arginine fingers were reported for several GTPase and AAA+ enzymes, including Myo9b-RhoGAP,[58] ClpB/Hsp104,[59] ClpA, ClpC, p97/VCP/Cdc48,[60] and NSF.[61] In these enzymes the available structural data suggests that the second arginine contacts the nucleotide sugar group, not the triphosphate chain. Our structural analysis over many NTP hydrolysis enzymes has identified that there is a well-defined structural region occupied by arginines (Figure 4a) around the γ-phosphate. This supports a consensus position of arginine fingers (either intrinsic[10] or "classical" from a distinct subunit) corresponding to R140, which is located at the same region with respect to the phosphate chain and the magnesium ion in a large number of nucleotide triphosphate converting enzymes.

Besides dUTPases, we identified only one other example among NTP pyrophosphatase or pyrophosphorylase enzymes that employs an Arg finger (Figure 5). This enzyme family performs various essential signaling and regulatory functions, and belongs to the eukaryotic ubiquitin activating (E1) enzymes[62] or to the analogous prokaryotic protein families, e.g., MoeB/MoaD[49] and Thif/ThiS[63] playing roles in the biosynthesis of molybdopterin, and thiamine, respectively. These Arg fingers, however, are not required for catalytic activity with the respective Ala mutations representing an active enzyme.[49] Note also that the currently available crystallographic structures do not have a Mg^{2+} ion cofactor bound, which may affect the positioning of the Arg residues (in particular for Thif, see Figure S7).

While the E1-like pyrophosphatase enzymes were reported to possess an Arg finger, they do not strictly fulfill all the required criteria, as the R/K or even R/A mutations do not affect the catalytic activity[49] (cf., criteria for Arg fingers above). This fact implies that E1-like pyrophosphatases in addition to dUTPases do not use their structurally analogous Arg fingers for the electrostatic stabilization of the TS. This contrasts the functional role of Arg fingers from ATPases and GTPases where these catalytically essential residues directly stabilize the TS. Nevertheless, as the R140 Arg finger structurally positions the P-loop-like motif of dUTPases, its presence is critical for ensuring optimal catalytic activity.

Evolutionary Development of dUTPases: Recruiting the Arg Finger and P-Loop-like Motif, and Emergence of Novel Functions. The β-fold dUTPase superfamily includes dCTP deaminases (DCD), bifunctional dCTP deaminase/ dUTPases (DCD-DUT), and dUTPases. These enzymes share similar overall fold and subunit assembly, yet the active site of DCD is assembled from only two and not from three subunits.[64] Among these enzymes, only monofunctional dUTPases employ the P-loop-like motif, even though DCD-DUT is also acting as a pyrophosphatase. It has been proposed that the Last Universal Common Ancestor (LUCA) of dUTPases could have possessed bifunctional DCD/DUT activity.[65] The monofunctional DUT enzymes were evolved later from this ancient promiscuous enzyme, according to the common proposed model of enzyme catalysis evolution.[66] An Arg or Lys, not coupled to a P-loop-like motif, is present in some DCD-DUTs at a sequential position corresponding to the fifth motif Arg (R140) of trimeric DUTs.[67] On the basis of the currently available structural information, however, these residues do not coordinate the cognate dCTP or dUTP nucleotides (e.g., PDB 2HXD). Instead, another conserved Lys coordinates the γ-phosphate of the nucleotide in DCD and DCD-DUT enzymes.[20] On the basis of these observations, we suggest that the acquisition of the P-loop-like motif might have enabled the antecedent arginine to adopt the function of an Arg finger. The novel Arg finger function together with the P-loop-like fifth motif probably contributed to their increased substrate specificity and dUTPase catalytic rate acceleration.[20]

Intriguingly, it has recently been revealed that some trimeric dUTPases possess additional functionality beyond their role in the nucleotide metabolism.[17,23,47] Moonlighting activity generally associated with viral transfection and expression regulation was reported for many viral dUTPases.[18,68−70] It has been previously suggested that several viruses acquired their *dut* gene from their host organism via horizontal gene transfer.[71] Although the underlying molecular mechanism is not yet fully deciphered, the acquired fifth motif together with other elements was suggested to play a role in the observed moonlighting function,[72] and within, the involvement of the Arg finger may also be envisaged. Nevertheless, in some cases the novel immunoregulatory function of viral dUTPases was presented in the absence of cognate dUTPase activity[73] and of the P-loop-like motif.[74,75]

CONCLUSION

We showed that the Arg from the fifth motif (R140) fulfills the criteria used to identify Arg fingers, therefore associating this conserved arginine of dUTPases with Arg fingers for the first time. We demonstrated that this highly conserved residue is required for full catalytic activity, promoting catalysis while being located on distinct subunit and exerting its effects on interprotomer interactions around the substrate. Functional differences between Arg fingers of dUTPases and of other Arg finger-employing P-loop NTPases may be explained on the basis of the different role of the cognate Arg finger in the catalytic mechanism. Notably, currently known enzymes with Arg fingers catalyze the cleavage between the $\beta-\gamma$ phosphate groups of the respective nucleotides, with the single exception of E1-like enzymes, for which, however, the conserved Arg fingers are not essential for catalysis. As the cleavage takes place between the $\alpha-\beta$ phosphates for pyrophosphatases, R140 being further from the cleavage site has a more pronounced structural role, whereas R64, another conserved Arg residue from a different subunit, can be directly involved in TS stabilization. Together, these findings support our hypothesis that the role of the Arg finger in dUTPases is to correctly position the P-loop residues.

The distinct structural composition of dUTPases renders them highly analogous to NTPases, and to GTPases in particular, albeit with a unique chemistry. It remains to be explored how these structural features are related to the recently identified moonlighting function that dUTPase enzymes play in viral signaling and how their evolutionary development took place.

ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b09012.

Additional figures, including ESI-MS spectra, steady-state kinetic measurements and ligand binding isotherms, active site structural observations, QM/MM energy profiles, structural comparison, and position of conserved Arg residues in enzymes with two metal-ion catalytic mechanism; additional tables, including molecular masses, crystallographic B factor analysis, and conservation of active site Arg positions in dUTPases (PDF)

AUTHOR INFORMATION

Corresponding Authors
*nagy.gergely@ttk.mta.hu
*vertessy@mail.bme.hu
*edina.rosta@kcl.ac.uk

Present Addresses
▽Laboratory of Molecular Physiology, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892, United States, and Astbury Centre for Structural Molecular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, United Kingdom.
○Szent István University, Ybl Miklós Faculty of Architecture and Civil Engineering, Budapest 1442, Hungary.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Hanson, P. I.; Whiteheart, S. W. Nat. Rev. Mol. Cell Biol. 2005, 6 (7), 519.
(2) Wittinghofer, A.; Vetter, I. R. Annu. Rev. Biochem. 2011, 80, 943.
(3) Glover, B. P.; McHenry, C. S. Cell 2001, 105 (7), 925.
(4) Jambrina, P. G.; Bohuszewicz, O.; Buchete, N.-V.; Kolch, W.; Rosta, E. Biochem. Soc. Trans. 2014, 42 (4), 784.
(5) Jambrina, P. G.; Rauch, N.; Pilkington, R.; Rybakova, K.; Nguyen, L. K.; Kholodenko, B. N.; Buchete, N.-V.; Kolch, W.; Rosta, E. Angew. Chem., Int. Ed. 2016, 55 (3), 983.
(6) Szabó, J. E.; Takács, E.; Merényi, G.; Vértessy, B. G.; Tóth, J. Sci. Rep. 2016, 6, 24219.
(7) Ahmadian, M. R.; Stege, P.; Scheffzek, K.; Wittinghofer, A. Nat. Struct. Biol. 1997, 4, 686−689.
(8) Hanson, P. I.; Whiteheart, S. W. Nat. Rev. Mol. Cell Biol. 2005, 6 (7), 519.
(9) Gasper, R.; Meyer, S.; Gotthardt, K.; Sirajuddin, M.; Wittinghofer, A. Nat. Rev. Mol. Cell Biol. 2009, 10 (6), 423.
(10) Schröter, G.; Mann, D.; Kötting, C.; Gerwert, K. J. Biol. Chem. 2015, 290 (28), 17085.
(11) Kagawa, R.; Montgomery, M. G.; Braig, K.; Leslie, A. G. W.; Walker, J. E. EMBO J. 2004, 23, 2734.
(12) Komoriya, Y.; Ariga, T.; Iino, R.; Imamura, H.; Okuno, D.; Noji, H. J. Biol. Chem. 2012, 287, 15134.
(13) Crampton, D. J.; Guo, S.; Johnson, D. E.; Richardson, C. C. Proc. Natl. Acad. Sci. U. S. A. 2004, 101, 4373.
(14) Danot, O.; Marquenet, E.; Vidal-Ingigliardi, D.; Richet, E. Structure 2009, 17 (2), 172.
(15) Yukawa, A.; Iino, R.; Watanabe, R.; Hayashi, S.; Noji, H. Biochemistry 2015, 54 (2), 472.
(16) Vértessy, B. G.; Tóth, J. Acc. Chem. Res. 2009, 42, 97.
(17) Tormo-Más, M. A.; Mir, I.; Shrestha, A.; Tallent, S. M.; Campoy, S.; Lasa, I.; Barbé, J.; Novick, R. P.; Christie, G. E.; Penadés, J. R. Nature 2010, 465 (7299), 779.
(18) Szabó, J. E.; Németh, V.; Papp-Kádár, V.; Nyíri, K.; Leveles, I.; Bendes, A. Á.; Zagyva, I.; Róna, G.; Pálinkás, H. L.; Besztercei, B.; Ozohanics, O.; Vékey, K.; Liliom, K.; Tóth, J.; Vértessy, B. G. Nucleic Acids Res. 2014, 42 (19), 11912.
(19) Mol, C. D.; Harris, J. M.; McIntosh, E. M.; Tainer, J. A. Structure 1996, 4 (9), 1077.

(20) Pécsi, I.; Szabó, J. E.; Adams, S. D.; Simon, I.; Sellers, J. R.; Vértessy, B. G.; Tóth, J. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (35), 14437.

(21) Vertessy, B. G.; Larsson, G.; Persson, T.; Bergman, A. C.; Persson, R.; Nyman, P. O. *FEBS Lett.* **1998**, *421*, 83.

(22) Nord, J.; Kiefer, M.; Adolph, H. W.; Zeppezauer, M. M.; Nyman, P. O. *FEBS Lett.* **2000**, *472* (2–3), 312.

(23) Németh-Pongrácz, V.; Barabás, O.; Fuxreiter, M.; Simon, I.; Pichová, I.; Rumlová, M.; Zábranská, H.; Svergun, D.; Petoukhov, M.; Harmat, V.; Klement, E.; Hunyadi-Gulyás, E.; Medzihradszky, K. F.; Kónya, E.; Vértessy, B. G. *Nucleic Acids Res.* **2007**, *35* (2), 495.

(24) Lopata, A.; Jambrina, P. G.; Sharma, P. K.; Brooks, B. R.; Toth, J.; Vertessy, B. G.; Rosta, E. *ACS Catal.* **2015**, *5* (6), 3225.

(25) Freeman, L.; Buisson, M.; Tarbouriech, N.; Van der Heyden, A.; Labbé, P.; Burmeister, W. P. *J. Biol. Chem.* **2009**, *284* (37), 25280.

(26) Shao, H.; Robek, M. D.; Threadgill, D. S.; Mankowski, L. S.; Cameron, C. E.; Fuller, F. J.; Payne, S. L. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1997**, *1339* (2), 181.

(27) Tchigvintsev, A.; Singer, A. U.; Flick, R.; Petit, P.; Brown, G.; Evdokimova, E.; Savchenko, A.; Yakunin, A. F. *Biochem. J.* **2011**, *437* (2), 243.

(28) Yang, W. *Nat. Struct. Mol. Biol.* **2008**, *15* (11), 1228.

(29) Varga, B.; Barabás, O.; Takács, E.; Nagy, N.; Nagy, P.; Vértessy, B. G. *Biochem. Biophys. Res. Commun.* **2008**, *373* (1), 8.

(30) Vertessy, B. G.; Persson, R.; Rosengren, A. M.; Zeppezauer, M.; Nyman, P. O. *Biochem. Biophys. Res. Commun.* **1996**, *219* (2), 294.

(31) Kabsch, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66* (2), 125.

(32) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66* (4), 486.

(33) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67* (4), 355.

(34) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67* (4), 235.

(35) *The PyMOL Molecular Graphics System, Version 1.8*; Schrödinger, LLC.

(36) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. *J. Comput. Chem.* **2008**, *29* (11), 1859.

(37) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L.; MacKerell, A. D.; Klauda, J. B.; Im, W. *J. Chem. Theory Comput.* **2016**, *12* (1), 405.

(38) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30* (10), 1545.

(39) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26* (16), 1781−1802.

(40) Huang, J.; MacKerell, A. D. *J. Comput. Chem.* **2013**, *34* (25), 2135.

(41) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98* (12), 10089.

(42) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, R. A., Jr; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F., III; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8* (27), 3172.

(43) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Mol. Phys.* **2015**, *113* (2), 184.

(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision E.01*; Gaussian, Inc.: Wallingford, CT, 2009.

(45) Woodcock, H. L.; Hodoscek, M.; Gilbert, A. T. B.; Gill, P. M. W.; Schaefer, H. F.; Brooks, B. R. *J. Comput. Chem.* **2007**, *28* (9), 1485.

(46) Takács, E.; Nagy, G.; Leveles, I.; Harmat, V.; Lopata, A.; Tóth, J.; Vértessy, B. G. *FEBS Lett.* **2010**, *584* (14), 3047.

(47) Pecsi, I.; Hirmondo, R.; Brown, A. C.; Lopata, A.; Parish, T.; Vertessy, B. G.; Tóth, J. *PLoS One* **2012**, *7* (5), e37461.

(48) Lopata, A.; Leveles, I.; Bendes, Á. Á.; Viskolcz, B.; Vértessy, B. G.; Jójárt, B.; Tóth, J. *J. Biol. Chem.* Under revision.

(49) Lake, M. W.; Wuebbens, M. M.; Rajagopalan, K. V.; Schindelin, H. *Nature* **2001**, *414* (6861), 325.

(50) Walker, J. E.; Saraste, M.; Runswick, M. J.; Gay, N. J. *EMBO J.* **1982**, *1*, 945.

(51) Vertessy, B. G. *Proteins: Struct., Funct., Genet.* **1997**, *28* (4), 568.

(52) Pecsi, I.; Leveles, I.; Harmat, V.; Vertessy, B. G.; Toth, J. *Nucleic Acids Res.* **2010**, *38* (20), 7179.

(53) Florián, J.; Warshel, A.; Goodman, M. F. *J. Phys. Chem. B* **2002**, *106* (22), 5754.

(54) Bork, P.; Koonin, E. V. *Proteins: Struct., Funct., Genet.* **1994**, *20*, 347.

(55) Law, A.; Boulanger, M. J. *J. Biol. Chem.* **2011**, *286*, 15577.

(56) Tesmer, J. J.; Klem, T. J.; Deras, M. L.; Davisson, V. J.; Smith, J. L. *Nat. Struct. Biol.* **1996**, *3*, 74.

(57) Yang, W.; Lee, J. Y.; Nowotny, M. *Mol. Cell* **2006**, *22* (1), 5.

(58) Yi, F.; Kong, R.; Ren, J.; Zhu, L.; Lou, J.; Wu, J. Y.; Feng, W. *J. Mol. Biol.* **2016**, *428* (15), 3043.

(59) Zeymer, C.; Fischer, S.; Reinstein, J. *J. Biol. Chem.* **2014**, *289* (47), 32965.

(60) Wang, Q.; Song, C.; Irizarry, L.; Dai, R.; Zhang, X.; Li, C.-C. H. *J. Biol. Chem.* **2005**, *280* (49), 40515.

(61) Zhao, C.; Matveeva, E. A.; Ren, Q.; Whiteheart, S. W. *J. Biol. Chem.* **2010**, *285* (1), 761.

(62) Schulman, B. A.; Harper, J. W. *Nat. Rev. Mol. Cell Biol.* **2009**, *10* (5), 319.

(63) Fujishiro, T.; Kahnt, J.; Ermler, U.; Shima, S. *Nat. Commun.* **2015**, *6*, 6895.

(64) Johansson, E.; Fanø, M.; Bynck, J. H.; Neuhard, J.; Larsen, S.; Sigurskjold, B. W.; Christensen, U.; Willemoës, M. *J. Biol. Chem.* **2005**, *280* (4), 3051.

(65) Iyer, L. M.; Aravind, L. *Proteins: Struct., Funct., Genet.* **2004**, *55* (4), 977.

(66) Khersonsky, O.; Roodveldt, C.; Tawfik, D. S. *Curr. Opin. Chem. Biol.* **2006**, *10* (5), 498.

(67) Helt, S. S.; Thymark, M.; Harris, P.; Aagaard, C.; Dietrich, J.; Larsen, S.; Willemoes, M. *J. Mol. Biol.* **2008**, *376* (2), 554.

(68) Weiss, R. S.; Lee, S. S.; Prasad, B. V.; Javier, R. T. *J. Virol.* **1997**, *71* (3), 1857.

(69) Davison, A. J.; Stow, N. D. *J. Virol.* **2005**, *79* (20), 12880.

(70) Tormo-Más, M. Á.; Donderis, J.; García-Caballer, M.; Alt, A.; Mir-Sanchis, I.; Marina, A.; Penadés, J. R. *Mol. Cell* **2013**, *49* (5), 947.

(71) McClure, M. A. *Curr. Protein Pept. Sci.* **2001**, *2* (4), 313.

(72) Maiques, E.; Quiles-Puchalt, N.; Donderis, J.; Ciges-Tomas, J. R.; Alite, C.; Bowring, J. Z.; Humphrey, S.; Penadés, J. R.; Marina, A. *Nucleic Acids Res.* **2016**, *44* (11), 5457.

(73) Madrid, A. S.; Ganem, D. *J. Virol.* **2012**, *86* (16), 8693.

(74) Voronin, N.; Herzig, E.; Hizi, A. *Retrovirology* **2014**, *11* (1), 60.

(75) Hizi, A.; Herzig, E. *Retrovirology* **2015**, *12* (1), 70.