

Structures of Yeast Apa2 Reveal Catalytic Insights into a Canonical Ap₄A Phosphorylase of the Histidine Triad Superfamily

Wen-Tao Hou¹, Wen-Zhe Li², Yuxing Chen¹, Yong-Liang Jiang¹ and Cong-Zhao Zhou¹

1 - Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei Anhui 230027, People's Republic of China

2 - Shanghai Key Laboratory for Signaling and Diseases, School of Life Science and Technology, Tongji University, 1239 Siping Road, Shanghai 200092, People's Republic of China

Correspondence to Yong-Liang Jiang and Cong-Zhao Zhou: jyl@mail.ustc.edu.cn; zcz@ustc.edu.cn

<http://dx.doi.org/10.1016/j.jmb.2013.04.018>

Edited by M. Guss

Abstract

The homeostasis of intracellular diadenosine 5',5'''-P¹,P⁴-tetrphosphate (Ap₄A) in the yeast *Saccharomyces cerevisiae* is maintained by two 60% sequence-identical paralogs of Ap₄A phosphorylases (Apa1 and Apa2). Enzymatic assays show that, compared to Apa1, Apa2 has a relatively higher phosphorylase activity towards Ap₃A (5',5'''-P¹,P³-tetrphosphate), Ap₄A, and Ap₅A (5',5'''-P¹,P⁵-tetrphosphate), and Ap₄A is the favorable substrate for both enzymes. To decipher the catalytic insights, we determined the crystal structures of Apa2 in the apo-, AMP-, and Ap₄A-complexed forms at 2.30, 2.80, and 2.70 Å resolution, respectively. Apa2 is an α/β protein with a core domain of a twisted eight-stranded antiparallel β-sheet flanked by several α-helices, similar to the galactose-1-phosphate uridylyltransferase (GalT) members of the histidine triad (HIT) superfamily. However, a unique auxiliary domain enables an individual Apa2 monomer to possess an intact substrate-binding cleft, which is distinct from previously reported dimeric GalT proteins. This cleft is perfectly complementary to the favorable substrate Ap₄A, the AMP and ATP moieties of which are perpendicular to each other, leaving the α-phosphate group exposed at the sharp turn against the catalytic residue His161. Structural comparisons combined with site-directed mutagenesis and activity assays enable us to define the key residues for catalysis. Furthermore, multiple-sequence alignment reveals that Apa2 and homologs represent canonical Ap₄A phosphorylases, which could be grouped as a unique branch in the GalT family.

© 2013 The Authors. Published by Elsevier Ltd. All rights reserved.

Introduction

Dinucleoside 5',5'''-polyphosphate molecules (Np_rN, where N stands for A, G, C or U) are a diverse group of nucleotide derivatives in all living organisms.¹ They are predominantly by-products of protein synthesis, specifically produced by aminoacyl-tRNA synthetases.² Their intracellular concentration is usually maintained at a level of 0.1–3 μM, which may be sharply elevated in response to stimuli such as heat shock, oxidative stress, nutrition shortage, and DNA damage.³ 5',5'''-P¹,P⁴-tetrphosphate (Ap₄A) is one of these derivatives that have been well investigated. It has been suggested that

Ap₄A acts as an intracellular “alarmone” in both prokaryotes⁴ and eukaryotes.⁵ The intracellular functions of Ap₄A are involved in apoptosis⁶ and DNA repair,⁷ whereas its extracellular roles are related to neurotransmitter and cardiac effects mediator.^{8,9} In bacteria, Ap₄A metabolism has also been implied in regulation of the stress response,¹⁰ pathogenesis,^{11–14} and antibiotic tolerance.¹⁵ Ap₄A metabolism also affects biofilm formation of *Pseudomonas fluorescens* via modulation of c-di-GMP-dependent pathways.¹⁶ In addition, a previous study has provided concrete evidence establishing Ap₄A as a second messenger in the regulation of gene expression.¹⁷

Specific degradation enzymes have been evolved to maintain the cellular homeostasis of these nucleotide derivatives. These enzymes are mainly grouped into three structurally distinct superfamilies:¹⁸ Nudix (nucleoside diphosphate linked to x), MPP (metallophosphoesterase), and HIT (histidine triad) superfamily. HIT proteins share a motif of H ϕ H ϕ H ϕ (where ϕ represents a hydrophobic amino acid) and act as nucleotide hydrolases and transferases. The middle histidine functions as the nucleophilic residue that attacks the α -phosphate group of nucleotides to generate a covalently linked enzyme–nucleotide intermediate.¹⁹ Based on the sequence, substrate specificity, structure, evolution, and catalytic mechanism, HIT proteins have been classified into four families: the Hint, Fhit, galactose-1-phosphate uridylyltransferase (GalT), and aprataxin family.¹⁹ The Hint family consists of adenosine 5'-monophosphoramidate hydrolases, whereas the Fhit family is composed of diadenosine polyphosphate hydrolases. Aprataxin leads a discrete family of proteins possessing both DNA/RNA binding and nucleotide hydrolase activities.²⁰ In contrast, the GalT family has a modified motif of HXHXQ ϕ . Members in the GalT family are further grouped into three branches: GalT, Ap₄A phosphorylase, and adenylylsulfate:phosphate adenylyltransferase. The key difference between Hint/Fhit of HIT hydrolase and GalT of HIT transferase is that the hydrolases can transfer the nucleotide directly to a water molecule, instead of requiring a second substrate to accept the histidine-bound nucleotide.¹⁹ In addition, Hint and Fhit proteins have two active sites and are typically homodimers of ~15 kDa/subunit, whereas GalT members share an imperfectly tandem-repeated polypeptide of ~30 kDa that retains a single active site. The HIT proteins have been maintained highly conserved throughout evolution, which implies that they fulfill basic, perhaps vital functions.²¹ To date, several crystal structures of HIT proteins have been solved,^{22–29} and two of which are from the GalT family.^{23,26} The crystal structure of *Mycobacterium tuberculosis* Rv2613c demonstrated a novel Ap₄A phosphorylase that shares a structure similar to Fhit Ap_rA hydrolases.²² The yeast *Saccharomyces cerevisiae* genome encodes two Ap₄A phosphorylases, Apa1 and Apa2, which are 60% sequence identical and grouped into the GalT family based on their primary sequences. They both catalyze the degradation of Ap₄A into ADP and ATP in the presence of inorganic phosphate, but at different catalytic activities.³⁰ With the disruption of either *apa1* or *apa2* alone, even both genes did not affect the cell viability; however, the absence of both genes leads to the dramatic increase of the intracellular Ap₄A concentration.³⁰ To date, no structural information of the Ap₄A phosphorylase in the GalT family is available. The crystal structures and biochemical analyses of yeast

Apa2 provided us insights into the substrate specificity and the catalytic mechanism of a canonical Ap₄A phosphorylase.

Results and Discussion

Apa2 represents a canonical Ap₄A phosphorylase

A previous study suggested that Apa1 and Apa2 are Ap₄A phosphorylases, which catalyze the phosphorolytic degradation of Ap₄A into ADP and ATP.³⁰ We thus checked the activity of our recombinant Apa1 and Apa2 towards Ap₄A using high-performance liquid chromatography (HPLC). We first screened the activity at the pH range from 4.5 to 9.0 and determined that both Apa1 and Apa2 had an optimal pH at 7.5. As a previous study suggested that HIT proteins require metal ions for the catalysis,²³ we systematically compared the activities of Apa2 in the presence of various metal ions. Compared to the activity without adding cations, the cations Co²⁺, Cd²⁺, Zn²⁺, Mg²⁺, Mn²⁺, and Ca²⁺ could increase the activity at varying ratios, with Mn²⁺ being the most significant. By contrast, the cations Ni²⁺, Fe²⁺, and Fe³⁺ had no significant effect (Fig. S1). A metal binding site could not be identified in structures of either apo-Apa2 or AMP- and Ap₄A-complexed Apa2, like previous reports of some HIT members.^{21,28} Thus, we suspected that catalysis should occur in the absence of a divalent metal. Indeed, Apa2 degrades Ap₄A, albeit at a slower rate, even in the presence of 5 mM ethylenediaminetetraacetic acid. Nucleotide polyphosphates usually exist in the cell as complexes with divalent cations. The effects of metal ions observed here may be resulted in part from the impact of divalent cations on nucleotide conformation.

The enzymatic assays of Apa1 and Apa2 towards Ap₄A were performed in the presence of 0.2 mM Mn²⁺ at pH 7.5. The K_m and k_{cat} values are $4.5 \pm 2.0 \mu\text{M}$ and $93.8 \pm 2.0 \text{ s}^{-1}$, respectively, for Apa2, and $39.3 \pm 2.0 \mu\text{M}$ and $81.2 \pm 12.7 \text{ s}^{-1}$ for Apa1, resulting to an activity (k_{cat}/K_m) of $21.0 \text{ s}^{-1} \mu\text{M}^{-1}$ for Apa2, 10-fold to that of Apa1 ($2.1 \text{ s}^{-1} \mu\text{M}^{-1}$) (Table 2). The Ap₄A phosphorylase activity of Apa2 is about 200-fold to that of *M. tuberculosis* Rv2613c, the structure of which is similar to that of Hint and Fhit members.²² It is worthy to notice that the K_m value of Apa2 is comparable to the physiological concentration of Ap₄A (0.1–3 μM). The cellular concentrations for Apa1 and Apa2 are approximately 1.12 and 0.10 μM , respectively, as calculated by 19,600 and 1770 molecules/cell divided by the cell volume of $29 \times 10^{-15} \text{ L}$.³¹ The lower intracellular concentration combined with a higher activity of Apa2 compared to that of Apa1 suggested that Apa2 and

Apa1 might function *in vivo* as a fine tuner and a robust cutter, respectively. When Ap₄A is accumulated to a high concentration upon certain stimuli, the high-abundance Apa1 could function rapidly as a robust cutter to reduce the concentration. In contrast, the low-abundance Apa2 works as a fine tuner to control the subtle change of Ap₄A. In addition, neither Apa1 nor Apa2 can catalyze the reverse reaction of Ap₄A synthesis using ADP and ATP as the substrates. These results clearly demonstrated that Apa1 and Apa2 are canonical Ap₄A phosphorylases.

The overall structure of Apa2 reveals the first monomeric GalT protein

The structure of apo-form Apa2 was determined at 2.30 Å resolution by single-wavelength anomalous diffraction phasing method using selenomethionine (SeMet)-substituted crystals. Each asymmetric unit contains two Apa2 molecules with a buried interface area of about 500 Å², which is not large enough to stabilize a homodimer. Size-exclusion chromatography with a Superdex 75 column gave an elution volume of 71.6 mL, corresponding to a calculated mass of approximately 37 kDa, which is comparable to that of the monomeric Apa2 (Fig. S2). The structures of Ap₄A- and AMP-complexed forms

were determined at 2.70 and 2.80 Å resolution, respectively, by molecular replacement using the apo-form structure as a search model. A molecule of Ap₄A or two molecules of AMP could be well defined at the active sites of each subunit in the corresponding complex structures. In the Apa2–AMP structure, a molecule of phosphate ion is bound to the loop_{β6–β7} and loop_{β10–η4} in molecule A, but not in molecule B, via two hydrogen bonds with H265 and H152. The final models exhibited good crystallographic and geometric statistics (Table 1), as evaluated with the programs MolProbity³² and PROCHECK.³³ In the apo-Apa2 model, residues Lys53–Glu65 and Lys134–Thr139 in molecule A and Lys53–Lys63 in molecule B are missing due to their poor electron density.

The overall structure of Apa2 resembles the GalT members of the HIT superfamily, which adopt an α/β fold (Fig. 1a and b). It can be divided into two domains: the core catalytic domain and the auxiliary domain. The core domain may be further dissected into two tandem HIT subdomains (HIT1 and HIT2). It folds into the central antiparallel open-faced β-pleated sheet formed by eight β-strands (β3–β10) that is cradled on both sides by six helical elements (α1–α4, η2, and η3). The conserved H¹⁵⁹KHLQIM¹⁶⁵ motif is located at β7 of the HIT1 subdomain. The auxiliary domain

Table 1. Crystal parameters, data collection, and structure refinement

	SeMet–Apa2	Apo–Apa2	Apa2–Ap ₄ A	Apa2–AMP
<i>Data collection</i>				
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>C</i> 2	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Unit cell				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	74.31, 74.01, 112.34	73.76, 74.46, 112.35	203.35, 52.78, 70.76	52.85, 70.33, 174.24
α, β, γ (°)	90.00	90.00	90.00, 99.92, 90.00	90.00
Resolution range (Å)	50.00–2.80 (2.90–2.80) ^a	47.49–2.30 (2.42–2.30)	50.00–2.70 (2.79–2.70)	50.00–2.80 (2.90–2.80)
Unique reflections	16,010 (1546)	27,391 (3415)	19,447 (1801)	15,784 (1601)
Completeness (%)	99.9 (100)	97.4 (86.1)	93.9 (96.7)	94.1 (99.3)
⟨ <i>I</i> /σ(<i>I</i>)⟩	24.9 (10.7)	10.5 (3.2)	7.4 (1.9)	8.6 (4.0)
<i>R</i> _{merge} ^b (%)	11.4 (30.0)	9.1 (32.6)	7.0 (34.8)	12.4 (57.0)
Average redundancy	13.5 (13.8)	5.7 (3.5)	2.2 (2.2)	4.8 (5.3)
<i>Structure refinement</i>				
Resolution range (Å)		44.84–2.30	42.07–2.70	43.56–2.80
<i>R</i> -factor ^c / <i>R</i> _{free} ^d (%)		20.3/27.8	18.1/26.1	22.4/26.1
Number of protein atoms		4870	4685	4559
Number of water atoms		223	52	55
RMSD ^e bond lengths (Å)		0.008	0.008	0.010
RMSD bond angles (°)		1.251	1.330	1.497
Mean <i>B</i> -factors (Å ²)		37.3	69.5	60.1
Ramachandran plot ^f (residues)				
Most favored (%)		97.3	94.5	95.1
Additional allowed (%)		2.5	5.0	4.9
Outliers (%)		0.2	0.5	0
PDB entry		4I5T	4I5V	4I5W

^a The values in parentheses refer to statistics in the highest bin.

^b $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl) \rangle$ is the mean value for its unique reflection; summations are over all reflections.

^c $R\text{-factor} = \frac{\sum_h |F_o(h) - F_c(h)|}{\sum_h F_o(h)}$, where F_o and F_c are the observed and calculated structure-factor amplitudes, respectively.

^d R_{free} was calculated with 5% of the data excluded from the refinement.

^e RMSD from ideal values.

^f Categories were defined by MolProbity.

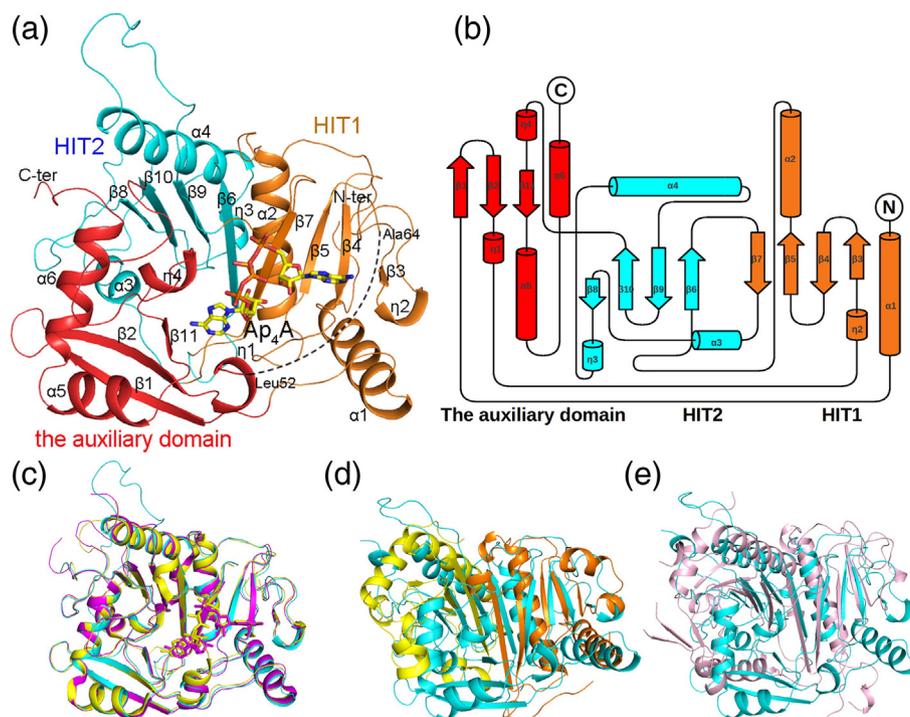


Fig. 1. Overall structure of apo-Apa2. (a) Cartoon representation of the apo-Apa2 (cyan and orange for the HIT1 and HIT2 of core domain, respectively, and red for the auxiliary domain). The secondary structural elements were labeled and the missing residues were depicted as broken lines. (b) A topology diagram of Apa2. (c) Structural comparison of apo-Apa2 (cyan) with Apa2–Ap₄A (yellow) and Apa2–AMP (purple). Structural comparison of apo-Apa2 with (d) the dimer of *M. tuberculosis* Rv2613c (yellow and orange for each subunit, PDB ID: 3ANO) and (e) *A. thaliana* ADP-glucose phosphorylase (pink, PDB ID: 1Z84).

consists of a curved three-stranded β -sheet (β 1, β 2, and β 11) with four helical elements (α 5, α 6, η 1, and η 4) packing around (Fig. 1a).

Structural comparisons of the apo-Apa2 against the Ap₄A- and AMP-complexed forms yielded a root-mean-square deviation (RMSD) of 0.525 Å and 0.379 Å, respectively, revealing no significant conformational changes of the overall structures (Fig. 1c). The loops of α 2– β 6, α 3– β 8, and α 4– β 9 located at the surface of Apa2 are relatively flexible and present diverse conformations. The overall *B*-factors of the three loops in apo-Apa2 are 71.6, 71.5, and 78.5, respectively, much higher than the average *B*-factor of 37.3 of the overall structure. Among the three loops, loop $_{\alpha$ 2– β 6 in Apa2–AMP and molecule A of apo-Apa2 and loop $_{\alpha$ 4– β 9 in Apa2–Ap₄A and Apa2–AMP were not modeled due to the poor electron density.

Dali protein structural similarity search† revealed that Apa2 resembles the members in the GalT family, despite a sequence identity of less than 15%. The closest member is the ADP-glucose phosphorylase from *Arabidopsis thaliana* [termed Ath, Protein Data Bank (PDB) ID: 2H39, Z-score = 11.1, RMSD = 4.0 Å over 196 C $^{\alpha}$ atoms]. Other members include *Escherichia coli* GalT (PDB ID: 1HXQ), *E.*

coli CDP-diacylglycerol pyrophosphatase (PDB ID: 2POF), followed by several more distinct homologs (Z-score = \leq 5.8) in the Hint and Fhit family. Superposition of Apa2 onto GalT or other HIT members revealed that the central core of the β -sheet and the surrounding two α -helices (α 2 and α 4) are similar, whereas the auxiliary domain varies a lot (Fig. 1d and e). The divergence of the auxiliary domain might contribute to the distinct substrate specificity among the HIT members.

Distinct from the previous dimeric or tetrameric HIT proteins of known structure,^{19,23,25,34} an individual Apa2 monomer possesses an intact active site. For example, dimerization of Ath²³ completes the assembly of active site by contributing residues *in trans* across the dimeric interface. In contrast, the auxiliary domain of Apa2 prevents the central β -sheet from forming dimeric interface but contributes to a side wall of the functional catalytic cleft.

Ap₄A is perfectly complementary to the interdomain cleft

To elucidate the structural details of Ap₄A-binding pattern, we co-crystallize Ap₄A with the H161A

mutant of Apa2 and solved the complex structure (termed Apa2–*Ap₄A*). The $F_o - F_c$ Fourier difference map is clear enough for us to fit a molecule of *Ap₄A* into the active site (Fig. 2a). The *Ap₄A*-binding cleft, with a buried interface of about 650 Å², is mainly formed by the central β-sheet (β4–β7) of the core domain, β11, helices η1, η4, and the connecting loops of the auxiliary domain (Fig. 2a). The electrostatics potential surface revealed that this deep cleft is mainly positively charged surrounding the phosphate groups but a little bit hydrophobic around the two adenosine bases of *Ap₄A* (Fig. 2b). Notably, the missing residues (Lys53–Glu65) from three structures of Apa2 were probably located across the *Ap₄A*-binding cleft, suggesting that these residues might be adjacent to the *Ap₄A* entrance tunnel and thus are relatively flexible. The shape and charge property of the cleft enable the precise positioning of *Ap₄A* in the active site. The AMP (the first adenine nucleotide and α-phosphate group) and ATP (the second adenine nucleotide, β-, γ-, and δ-phosphate groups) moieties of *Ap₄A* are perpendicular to each other, leaving the α-phosphate group at the sharp turn against the catalytic residue of the middle His161 in the HIT motif (Fig. 2c). *Ap₄A* is mainly stabilized by van der Waals interactions and hydrogen bonds. The AMP moiety is aligned approximately perpendicular to the direction of the central β-strands (β4–β7). Its adenosine base is stabilized by hydrophobic interactions with residues Pro67, Phe68, Phe94, and Leu102. The hydroxyl groups of the ribose make two hydrogen bonds with Asn92 and α-phosphate group is further stabilized by three hydrogen bonds with Asn148, Ser155, and Ser156 (Fig. 2c). In contrast, the ATP moiety is positioned almost parallel to the direction of the central β-strands. The adenosine base is fixed by hydrophobic interactions with Met284 and Leu286, and hydrogen-bonds with Thr179 and the backbone Met284. The ribose makes two hydrogen bonds with Asn92 and Gln163. The three phosphate groups form several hydrogen bonds with residues Lys53, Gly154, Ser156, Gln163, Asn277, and Lys288 (Fig. 2c). In the Apa2–AMP structure, two AMP molecules bind to the corresponding sites of AMP and ATP moieties of *Ap₄A* with a similar conformation (Fig. 2d). The phosphate group of the first AMP molecule is hydrogen-bonded by His161 and Gln163. Active-site comparisons between apo-Apa2 and Apa2–*Ap₄A* or Apa2–AMP complex revealed that all active-site residues could be well superimposed, suggesting that there is no induced fit of the active site upon the binding of *Ap₄A* or AMP (Fig. 2d). In contrast, active-site comparisons between Apa2–*Ap₄A* and Ath (PDB ID: 2H39) or Rv2613c (PDB ID: 3ANO) showed large divergence of the active sites (Fig. 2e and f). Despite the capability of AMP moieties and corresponding AMP-binding residues of both Apa2

and Ath to be superimposed, the ATP moiety of Apa2 and the counterpart glucose-6P of Ath differ a lot (Fig. 2e). Notably, three residues, G321, F325, and E326, stabilizing glucose-6P are contributed by the symmetric subunit in Ath, whereas the corresponding residues, M284, L286, and K288, in Apa2 are from the auxiliary domain. Comparison of the active sites of Apa2–*Ap₄A* and Rv2613c in complex with phosphate ion revealed that they adopt a similar active-site pocket for accommodating *Ap₄A*, despite that residues N277, T279, M284, L286, and K288 in the auxiliary domain of Apa2 correspond to residues N165, I167, S173, V175, and P177, respectively, which are contributed by the symmetric dimer of tetrameric Rv2613c (Fig. 2f). These results demonstrated that the members of the HIT superfamily share a conserved part for accommodating the mononucleotide moiety and differ from each other with the divergent part for recognizing diverse substrates.

Multiple-sequence alignment revealed that most *Ap₄A*-binding residues are conserved, except for Phe68 of Apa2, which corresponds to Leu67 of Apa1. This suggests that Apa1 and Apa2 share a similar *Ap₄A*-binding pattern and catalytic mechanism. To probe the structural determinant of a 10-fold higher activity of Apa2 compared to Apa1, the F68L and F68A mutants of Apa2 were subjected to enzymatic assays. The F68L mutant had a K_m of $8.0 \pm 1.1 \mu\text{M}$ and a k_{cat} of $48.5 \pm 1.8 \text{ s}^{-1}$, resulting to an activity of $6.1 \text{ s}^{-1} \mu\text{M}^{-1}$, which is one-fourth to that of Apa2 or 3-fold to that of Apa1. In contrast, the F68A mutant had an activity of $0.3 \text{ s}^{-1} \mu\text{M}^{-1}$, which is only about 1/70 and 1/7 compared to that of the wild-type Apa2 and Apa1, respectively (Table 2). The results suggested that a hydrophobic residue phenylalanine or leucine at position 68 of Apa2, which stacks the AMP moiety of *Ap₄A*, is important for the activity.

A putative catalytic mechanism of Apa2 towards *Ap₄A*

A previous report suggested that Apa1 and Apa2 could also degrade 5',5'''-*P*¹,*P*³-tetrphosphate (*Ap₃A*) and 5',5'''-*P*¹,*P*⁵-tetrphosphate (*Ap₅A*), in addition to *Ap₄A*.³⁰ We thus tested the enzymatic activities of Apa1 and Apa2 towards *Ap₃A* and *Ap₅A*.

Table 2. The kinetic parameters of Apa1, Apa2, and their mutants towards *Ap₄A*

Enzyme		K_m (μM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ μM ⁻¹)
Apa2	Wild type	4.5 ± 2.0	93.8 ± 2.0	21.0
	Q163H	3.2 ± 1.0	12.6 ± 0.6	4.0
	F68L	8.0 ± 1.1	48.5 ± 1.8	6.1
	F68A	40.0 ± 7.0	11.7 ± 0.5	0.3
	H161A	ND	ND	ND
	Wild type	39.3 ± 2.0	81.2 ± 12.7	2.1

Table 3. The relative activity of Apa1 and Apa2 towards *Ap₃A*, *Ap₄A*, and *Ap₅A*

	Apa2 (%)	Apa1 (%)
<i>Ap₃A</i>	8.3	<5
<i>Ap₄A</i>	100	13.8
<i>Ap₅A</i>	32.7	7.2

The enzymes and all substrates were present at a concentration of 10 nM and 0.5 mM, respectively. The reaction time is 15 min.

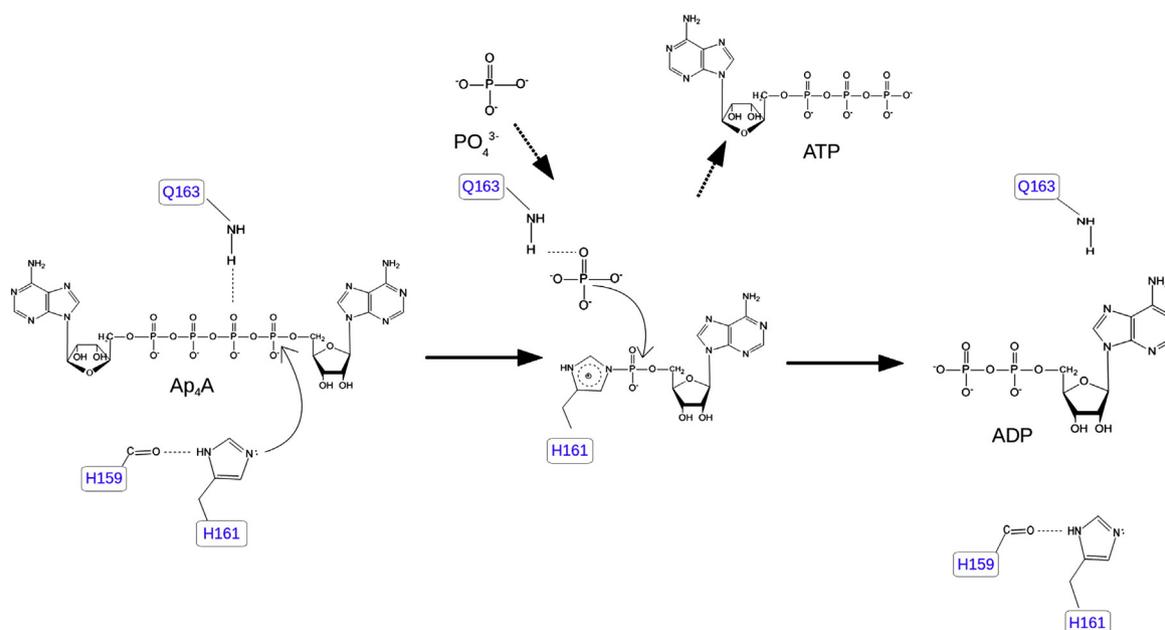
Both Apa1 and Apa2 could perform phosphorolytic degradation of *Ap₃A* and *Ap₅A*, producing two ADP molecules for *Ap₃A* and one molecule of *Ap₄* and ADP for *Ap₅A*. The relative activities of Apa2 towards *Ap₃A* and *Ap₅A* are 8.3% and 32.7%, respectively, to the activity towards *Ap₄A* (Table 3). As for Apa1, the relative activities are <5% (*Ap₃A*), 13.8% (*Ap₄A*), and 7.2% (*Ap₅A*), respectively (Table 3). The results clearly demonstrated that both Apa1 and Apa2 showed a much higher activity towards *Ap₄A*.

In the Apa2–*Ap₄A* complex, *Ap₄A* is buried deeply in the active-site cleft, with an interface area of about 650 Å². The total surface area of *Ap₄A* is about 640 Å², which is in good agreement with the active-site cleft. In addition, in the Apa2–AMP complex, two AMP molecules occupy the position of the corresponding ATP and AMP moieties of *Ap₄A*, leaving a distance of 8.0 Å between the two α-phosphate groups (Fig. 2d). This distance could perfectly accommodate two additional phosphate groups. Thus, *Ap₄A* is the most favorable substrate for both Apa1 and Apa2. *Ap₅A* binds to the active site perhaps with its five phosphate groups in a tighter

conformation, whereas *Ap₃A* with only three phosphate groups could not fully occupy the pocket.

The reactions catalyzed by Apa2 and other GalT members are chemically similar, both catalyzing nucleotidyl transfer with nucleotide derivatives as donor substrates. Given the sequence and structure homology, the chemical and kinetic mechanisms might be analogous. As previously proposed, a two-step double-displacement mechanism for the HIT superfamily²⁵ involves the transfer of the AMP nucleotide to the active site of the enzyme through nucleophilic attack, forming a covalent enzyme–AMP intermediate, followed by AMP hydrolysis or transfer of AMP to the acceptors. The present structures combined with the previous report²³ enabled us to propose a similar catalytic mechanism of Apa2 (Fig. 3).

The apo-Apa2 structure shows that the middle His161 in the modified HIT motif (H¹⁵⁹KHLQIM¹⁶⁵) is stabilized by the carbonyl oxygen atom of His159, suggesting the proper activation of these residues as nucleophiles for an inline attack on the α-phosphate group of *Ap₄A* as it enters the binding cleft. As His161 was mutated to Ala in the Apa2–*Ap₄A* complex, we superimposed the apo-Apa2 onto Apa2–*Ap₄A* complex and found that His161 is aligned in an optimal orientation to attack the α-phosphate group, at a distance of 2.9 Å counting from the N^ε of His161. Moreover, one α-phosphate group is 3.0 Å from the N^ε of His161 in the Apa2–AMP complex. Mutation of His161 to alanine almost completely abolished the activity towards *Ap₄A*, suggesting the indispensable role of His161 for the catalysis (Table 2). Our structural and biochemical

**Fig. 3.** A scheme showing the catalytic mechanism of Apa2.

studies support a double-displacement mechanism, the first step of which involves nucleophilic attack of Ap_4A by His161 to generate a covalent enzyme–AMP intermediate. In the second step, the third HIT residue Gln163 is suggested to activate the acceptor phosphate for general acid–base catalysis to release the intermediate AMP and produce ADP. In the $Apa2$ – Ap_4A structure, Gln163 contributes two hydrogen bonds with β -phosphate and the ribose of the ATP moiety (Fig. 2c). In addition, Gln163 in the $Apa2$ –AMP complex also makes a hydrogen bond with AMP phosphate oxygen (Fig. 2d). These findings suggested that Gln163 is favorably positioned to activate the acceptor phosphate. Furthermore, Gln163 may also play a critical role for orienting Ap_4A in a suitable conformation and further stabilizing the intermediate of the first step. The key difference between GalT transferases and Fhit/Hint hydrolases is that GalT has a modified HIT motif with a glutamine residue corresponding to the third histidine in Fhit/Hint. To see whether reverse of this modification is sufficient to restore the hydrolytic activity, we constructed a Q163H mutant and detected its activity. Compared to the wild-type $Apa2$, the Q163H mutant had transferase activity towards Ap_4A in the presence of phosphate, with a comparable K_m value of $3.2 \pm 1.0 \mu M$ but a much lower k_{cat} value of $12.6 \pm 0.6 s^{-1}$ (Table 2). However, the Q163H mutant did not show any hydrolytic activity towards Ap_4A without the addition of phosphate. A comparable K_m value suggested that the substitution of glutamine with histidine did not alter the Ap_4A binding affinity. Our results suggest that distant members in the HIT superfamily share a common catalytic mechanism.

$Apa2$ defines a unique branch of the GalT family

To date, the representative structures from two branches, GalT and ADP-glucose phosphorylase, of the GalT family are available.^{23,26,35} However, there is no three-dimensional structure information from the Ap_4A phosphorylase branch, which is only found in prokaryotes or unicellular eukaryotes, such as *Arthrospira* and *S. cerevisiae*. To address the evolutionary hints of the three branches in the GalT family, we performed a phylogenetic analysis with members from all branches to gain a more detailed insight (Fig. 4a). The uridylyltransferase branch exists in all species whereas the other two branches exist almost exclusively in prokaryotes or unicellular eukaryotes. In addition, the uridylyltransferase and adenylyltransferase branches are closely related, whereas the 60% sequence-identical $Apa1$ and $Apa2$ form a more discrete cluster (Fig. 4a).

Multiple-sequence alignment of $Apa2$ branch revealed that Ap_4A -binding residues as well as the HXHXQ $\phi\phi$ motif are conserved (Fig. 4b). It is suggested that these orthologs adopt a substrate-

binding pattern and catalytic mechanism similar to $Apa2$. Despite the shared HXHXQ $\phi\phi$ motif, $Apa2$ shares a sequence identity of less than 20% with the members from the other two branches, which attribute the diverse substrate specificity to the GalT members. Among the three branches, the HXHXQ $\phi\phi$ motif is strictly conserved. However, a number of conserved Ap_4A -binding residues, such as Pro67, Phe68, Asn92, Asn148, Gly154, Ser156, and Asn277, distinguish Ap_4A phosphorylases from members of the other two branches. Considering the highly conserved HIT motif and a similar overall structure, we propose that HIT members are evolved from a same ancestor. Afterwards, $Apa2$ and homologs, only found in prokaryotes or unicellular eukaryotes, evolved independently to form a discrete branch to handle the substrate Ap_4A . Particularly in the yeast *S. cerevisiae*, gene duplication resulted in the coexistence of two 60% sequence-identical paralogs, $Apa1$ and $Apa2$, which function *in vivo* as a robust cutter and a fine tuner, respectively, to maintain the homeostasis of intracellular Ap_4A .

Materials and Methods

Cloning, expression, and purification of $Apa2$ and mutants

Full-length *apa2* was PCR amplified from the genome of *S. cerevisiae* ATCC 204508/S288c and cloned into a pET-28a-derived vector, with an N-terminal hexahistidine tag. The recombinant plasmid was transformed into *E. coli* BL21 (DE3) cultured in $2 \times$ YT medium (5 g of NaCl, 16 g of Bacto-Tryptone, and 10 g of yeast extract per liter). The expression of the proteins was induced with 0.2 mM IPTG when OD_{600} reached 0.6. After being cultured for another 20 h at 16 °C, cells were harvested and collected by centrifugation at 4000g for 15 min, and resuspended in 40 mL of lysis buffer (20 mM Hepes, pH 8.0, and 100 mM NaCl). After sonication for 2.5 min and centrifugation at 12,000g for 30 min, the target proteins in the supernatant were collected and loaded onto a Ni-NTA column (GE Healthcare) that was pre-equilibrated with the binding buffer (20 mM Hepes, pH 8.0, and 100 mM NaCl). The target proteins were eluted with 300 mM imidazole and further applied to a Superdex 75 column (GE Healthcare) equilibrated with the binding buffer. The target proteins were collected and concentrated to 18 mg/mL for crystallization. Protein purity was assessed by electrophoresis and the samples were stored at –80 °C for further use.

The SeMet-labeled $Apa2$ protein was expressed in *E. coli* strain B834 (DE3) (Novagen). The cells were cultured in SeMet medium (M9 medium with 25 mg/L SeMet and other essential amino acids at 50 mg/L) to an OD_{600} of approximately 0.6. The following steps in protein expression and purification were the same as those for the native protein.

QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) was used to perform site-directed mutagenesis of $Apa2$. The expression, purification, and storage of the mutant proteins resembled that of the wild-type protein.

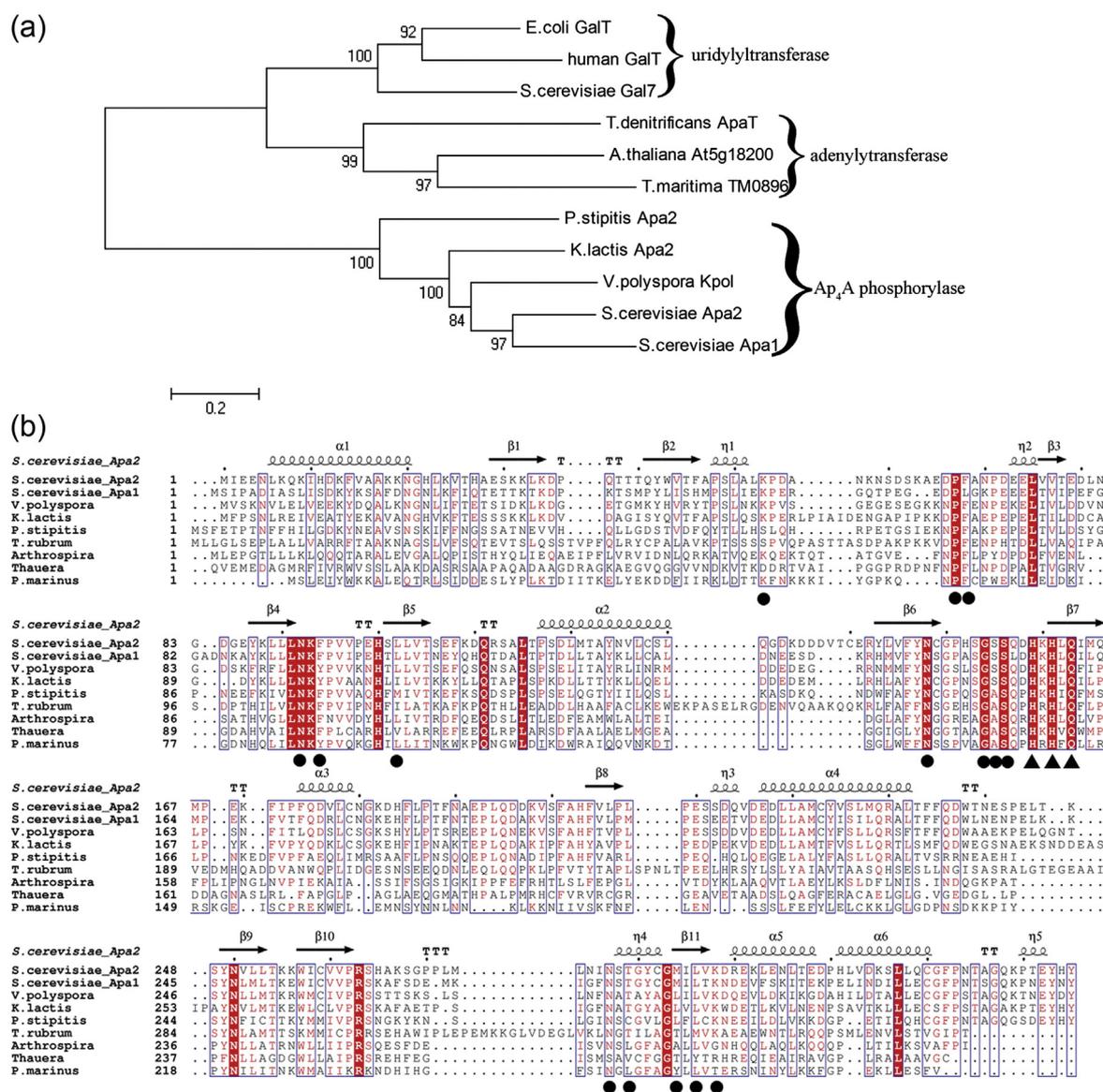


Fig. 4. Sequence conservation among Apa2 and homologs. (a) Phylogenetic relationships of the GalT family of proteins, with representatives from the three branches. The protein sequences from each branch (indicated by the bracket) were aligned by ClustalW and then used to build a neighbor joining tree. The bootstrap test was applied with 1000 replicates, and the values are indicated on the top at each node. (b) Multiple-sequence alignment of Apa2 and homologs. The secondary structural elements of Apa2 were labeled on the top. The Ap_4A -binding residues were labeled with black dots and the shared HXHXQFF motif was marked with black triangles. Colors are chosen according to rules of ESPript (<http://esprict.ibcp.fr/ESPript/ESPript/>): A blue frame represents a similarity across groups; a red character indicates similarity in a group; and a red box, white character demonstrates strict identity. All sequences were downloaded from the Swiss-Prot database (<http://www.uniprot.org/uniprot/>). The sequences are (Swiss-Prot accession numbers are in parentheses) *E. coli* GalT (P09148), human GalT (P07902), *S. cerevisiae* Gal7 (P08431), *Thiobacillus denitrificans* ApaT (Q9LA72), *A. thaliana* At5g18200 (Q9FK51), *Thermotoga maritima* TM0896 (Q9WZZ8), *S. cerevisiae* Apa2 (P22108), *S. cerevisiae* Apa1 (P16550), *Vanderwaltozyma polyspora* Kpol (A7TFD5), *Kluyveromyces lactis* Apa2 (P49348), *Pichia stipitis* Apa2 (A3LXQ2), *Trichophyton rubrum* bis(5'-nucleosyl)-tetraphosphatase (F2SIU7), *Arthrospira* Apa2 (K1WAN5), *Thauera* Apa2 (C4KBP8), and *Perkinsus marinus* ATP adenylyltransferase-like protein (Q31AV9).

Crystallization, data collection, and processing

Native and SeMet-substituted Apa2 were both concentrated to 18 mg/mL for crystallization. Crystals of apo-Apa2 were grown at 16 °C by the sitting drop vapor diffusion

method, with the initial condition of mixing 1 μ L of protein solution with an equal volume of the reservoir solution (18% polyethylene glycol 2K and 0.1 M 4-morpholineethanesulfonic acid, pH 6.5). The Apa2-AMP complex crystals were grown in the same manner as apo-Apa2 by the

addition of AMP and K₃PO₄ to a final concentration of 100 mM and 10 mM, respectively. The Apa2–Ap₄A complex proteins were obtained by mixing 18 mg/mL H161A mutant of Apa2 with Ap₄A (Sigma) to a final concentration of 10 mM. Crystals were grown at 16 °C using the hanging drop vapor diffusion method, with the initial condition of mixing 3 µL of protein solution with 1 µL of the reservoir solution (23% polyethylene glycol 2K, 0.1 M Tris–HCl, pH 8.0, and 20% glycerol). The crystals were transferred to cryoprotectant (reservoir solution supplemented with 30% glycerol) and flash-cooled with liquid nitrogen. All the data were collected at 100 K in a liquid nitrogen stream using beamline BL17U with a Q315r CCD (ADSC, MARresearch, Germany) at the Shanghai Synchrotron Radiation Facility. All the data were integrated and scaled with the program HKL2000.³⁶

Structure determination and refinement

The crystal structure of apo-Apa2 was determined using the single-wavelength anomalous diffraction phasing³⁷ method from a single SeMet-substituted protein crystal to a maximum resolution of 2.80 Å. The SHELXD program³⁸ implemented in IPCAS was used to locate the selenium atoms, and the phase was calculated by OASIS³⁹ and further improved with the programs RESOLVE and Buccaneer.^{40–42} Electron density maps showed clear features of secondary structural elements. Automatic model building was carried out using Autobuild in PHENIX.⁴³ The initial model was refined in REFMAC5⁴⁴ and Phenix.refine and rebuilt interactively using the program Coot.⁴⁵ The model was used as the search model against the native data of 2.3 Å by molecular replacement using the Molrep program as part of the CCP4i⁴⁶ program suite. The complex structures of Apa2–Ap₄A and Apa2–AMP were determined by molecular replacement using the apo-Apa2 as the search model. The final models were evaluated with the programs MolProbity³² and PROCHECK.³³ Crystallographic parameters are listed in Table 1. All structure figures were prepared with PyMOL.⁴⁷

Enzymatic activity assays

Protein samples for enzymatic activity assays were collected without concentration and were stored at –80 °C with the addition of 50% glycerol. The enzyme kinetic parameters of native Apa2 and its mutants were measured using Ap₃A, Ap₄A, and Ap₅A (Sigma) as the substrates. All the assays were performed at 37 °C in the buffer containing 50 mM Tris–HCl, pH 7.5, 10 mM K₃PO₄, and 0.2 mM MnCl₂. The pH and the addition of metal ion were optimized by preliminary experiment. The reactions were initiated by the addition of protein samples and terminated by heating at 100 °C for 5 min. The phosphorolytic degradation of substrates was measured by HPLC (Agilent 1200 series). Apa2 and its mutants were incubated with the substrate ranging from 3 to 100 µM in a volume of 25 µL. An assay mixture without any protein added was used as a control. ATP and ADP standards were quantified by HPLC using a series of concentrations ranging from 0.005 to 0.1 mM. The column (Eclipse XDB-C18 column, 4.6 × 150 mm, Agilent) was equilibrated using the buffer 20 mM NH₄H₂PO₄, pH 6.2. The samples were injected in a volume of 10 µL and were separated at a flow rate of

1 mL/min. The products were monitored with the absorption at 254 nm. The parameters K_m and k_{cat} were calculated by nonlinear fitting to the Michaelis–Menten equation using the program Origin 7.5. Three independent assays were performed to calculate the means and standard deviations for all K_m and k_{cat} values.

Accession numbers

The atomic coordinates and structure factors have been deposited into the PDB under the following accession codes: 4I5T (apo-Apa2), 4I5V (Apa2 in complex with Ap₄A), and 4I5W (Apa2 in complex with AMP).

Acknowledgements

We thank the staff at the Shanghai Synchrotron Radiation Facility for the data collection. We are grateful to all the developers of the CCP4 Suite, ESPript, MolProbity, and PyMOL. This work was supported by the Ministry of Science and Technology of China (Project Nos. 2012CB911000 and 2009CB918800).

Supplementary Data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jmb.2013.04.018>

Received 6 March 2013;

Received in revised form 17 April 2013;

Accepted 19 April 2013

Available online 26 April 2013

Keywords:

crystal structure;
substrate specificity;
Ap₄A phosphorylase;
enzymatic activity;
HIT superfamily

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

† http://ekhidna.biocenter.helsinki.fi/dali_server/

Abbreviations used:

Ap₄A, 5',5'''-P¹,P⁴-tetrphosphate; Ap₃A, 5',5'''-P¹,P³-tetrphosphate; Ap₅A, 5',5'''-P¹,P⁵-tetrphosphate; GalT, galactose-1-phosphate uridylyltransferase; HIT, histidine triad; HPLC, high-performance liquid chromatography; PDB, Protein Data Bank; SeMet, selenomethionine.

References

- Garrison, P. & Barnes, L. (1992). In *Ap₄A and Other Dinucleoside Polyphosphates* (McLennan, A. G., ed.), pp. 29–61, CRC Press, Boca Raton, FL.
- Goerlich, O., Foeckler, R. & Holler, E. (1982). Mechanism of synthesis of adenosine(5')tetraphospho(5')adenosine (AppppA) by aminoacyl-tRNA synthetases. *Eur. J. Biochem.* **126**, 135–142.
- Kisselev, L. L., Justesen, J., Wolfson, A. D. & Frolova, L. Y. (1998). Diadenosine oligophosphates (Ap(n)A), a novel class of signalling molecules? *FEBS Lett.* **427**, 157–163.
- Lee, P. C., Bochner, B. R. & Ames, B. N. (1983). AppppA, heat-shock stress, and cell oxidation. *Proc. Natl. Acad. Sci. USA*, **80**, 7496–7500.
- Varshavsky, A. (1983). Diadenosine 5', 5'''-P₁, P₄-tetraphosphate: a pleiotropically acting alarmone? *Cell*, **34**, 711–712.
- Vartanian, A., Prudovsky, I., Suzuki, H., Dal Pra, I. & Kisselev, L. (1997). Opposite effects of cell differentiation and apoptosis on Ap₃A/Ap₄A ratio in human cell cultures. *FEBS Lett.* **415**, 160–162.
- McLennan, A. G. (2000). Dinucleoside polyphosphates—friend or foe? *Pharmacol. Ther.* **87**, 73–89.
- Stavrou, B. M. (2003). Diadenosine polyphosphates: postulated mechanisms mediating the cardiac effects. *Curr. Med. Chem. Cardiovasc. Hematol. Agents*, **1**, 151–169.
- Pintor, J., Diaz-Hernandez, M., Gualix, J., Gomez-Villafuertes, R., Hernando, F. & Miras-Portugal, M. T. (2000). Diadenosine polyphosphate receptors: from rat and guinea-pig brain to human nervous system. *Pharmacol. Ther.* **87**, 103–115.
- Bochner, B. R., Lee, P. C., Wilson, S. W., Cutler, C. W. & Ames, B. N. (1984). AppppA and related adenylylated nucleotides are synthesized as a consequence of oxidation stress. *Cell*, **37**, 225–232.
- Mitchell, S. J. & Minnick, M. F. (1995). Characterization of a two-gene locus from *Bartonella bacilliformis* associated with the ability to invade human erythrocytes. *Infect. Immun.* **63**, 1552–1562.
- Cartwright, J. L., Britton, P., Minnick, M. F. & McLennan, A. G. (1999). The *lalA* invasion gene of *Bartonella bacilliformis* encodes a (de)nucleoside polyphosphate hydrolase of the MutT motif family and has homologs in other invasive bacteria. *Biochem. Biophys. Res. Commun.* **256**, 474–479.
- Badger, J. L., Wass, C. A. & Kim, K. S. (2000). Identification of *Escherichia coli* K1 genes contributing to human brain microvascular endothelial cell invasion by differential fluorescence induction. *Mol. Microbiol.* **36**, 174–182.
- Ismail, T. M., Hart, C. A. & McLennan, A. G. (2003). Regulation of dinucleoside polyphosphate pools by the YgdP and ApaH hydrolases is essential for the ability of *Salmonella enterica* serovar typhimurium to invade cultured mammalian cells. *J. Biol. Chem.* **278**, 32602–32607.
- Hansen, S., Lewis, K. & Vulic, M. (2008). Role of global regulators and nucleotide metabolism in antibiotic tolerance in *Escherichia coli*. *Antimicrob. Agents Chemother.* **52**, 2718–2726.
- Monds, R. D., Newell, P. D., Wagner, J. C., Schwartzman, J. A., Lu, W., Rabinowitz, J. D. & O'Toole, G. A. (2010). Diadenosine tetraphosphate (Ap₄A) metabolism impacts biofilm formation by *Pseudomonas fluorescens* via modulation of c-di-GMP-dependent pathways. *J. Bacteriol.* **192**, 3011–3023.
- Carmi-Levy, I., Yannay-Cohen, N., Kay, G., Razin, E. & Nechushtan, H. (2008). Diadenosine tetraphosphate hydrolase is part of the transcriptional regulation network in immunologically activated mast cells. *Mol. Cell. Biol.* **28**, 5777–5784.
- Guranowski, A. (2000). Specific and nonspecific enzymes involved in the catabolism of mononucleoside and dinucleoside polyphosphates. *Pharmacol. Ther.* **87**, 117–139.
- Brenner, C. (2002). Hint, Fhit, and GalT: function, structure, evolution, and mechanism of three branches of the histidine triad superfamily of nucleotide hydrolases and transferases. *Biochemistry*, **41**, 9003–9014.
- Kijas, A. W., Harris, J. L., Harris, J. M. & Lavin, F. M. (2006). Aprataxin forms a discrete branch in the HIT (histidine triad) superfamily of proteins with both DNA/RNA binding and nucleotide hydrolase activities. *J. Biol. Chem.* **281**, 13939–13948.
- Seraphin, B. (1992). The HIT protein family: a new family of proteins present in prokaryotes, yeast and mammals. *DNA Seq.* **3**, 177–179.
- Mori, S., Shibayama, K., Wachino, J. & Arakawa, Y. (2011). Structural insights into the novel diadenosine 5',5'''-P(1), P-tetraphosphate phosphorylase from *Mycobacterium tuberculosis* H37Rv. *J. Mol. Biol.* **410**, 93–104.
- McCoy, J. G., Arabshahi, A., Bitto, E., Bingman, C. A., Ruzicka, F. J., Frey, P. A. & Phillips, G. N., Jr. (2006). Structure and mechanism of an ADP-glucose phosphorylase from *Arabidopsis thaliana*. *Biochemistry*, **45**, 3154–3162.
- Krakowiak, A., Pace, H. C., Blackburn, G. M., Adams, M., Mekhafia, A., Kaczmarek, R. *et al.* (2004). Biochemical, crystallographic, and mutagenic characterization of hint, the AMP-lysine hydrolase, with novel substrates and inhibitors. *J. Biol. Chem.* **279**, 18711–18716.
- Lima, C. D., Klein, M. G. & Hendrickson, W. A. (1997). Structure-based analysis of catalysis and substrate definition in the HIT protein family. *Science*, **278**, 286–290.
- Wedekind, J. E., Frey, P. A. & Rayment, I. (1995). Three-dimensional structure of galactose-1-phosphate uridylyltransferase from *Escherichia coli* at 1.8 Å resolution. *Biochemistry*, **34**, 11049–11061.
- Brenner, C., Garrison, P., Gilmour, J., Peisach, D., Ringe, D., Petsko, G. A. & Lowenstein, J. M. (1997). Crystal structures of HINT demonstrate that histidine triad proteins are GalT-related nucleotide-binding proteins. *Nat. Struct. Biol.* **4**, 231–238.
- Pace, H. C., Garrison, P. N., Robinson, A. K., Barnes, L. D., Draganescu, A., Rosler, A. *et al.* (1998). Genetic, biochemical, and crystallographic characterization of Fhit-substrate complexes as the active signaling form of Fhit. *Proc. Natl. Acad. Sci. USA*, **95**, 5484–5489.

29. Tumbale, P., Appel, C. D., Kraehenbuehl, R., Robertson, P. D., Williams, J. S., Krahn, J. *et al.* (2011). Structure of an aprataxin–DNA complex with insights into AOA1 neurodegenerative disease. *Nat. Struct. Mol. Biol.* **18**, 1189–1195.
30. Plateau, P., Fromant, M., Schmitter, J. M. & Blanquet, S. (1990). Catabolism of bis(5′-nucleosidyl) tetraphosphates in *Saccharomyces cerevisiae*. *J. Bacteriol.* **172**, 6892–6899.
31. Ghaemmaghani, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephoure, N. *et al.* (2003). Global analysis of protein expression in yeast. *Nature*, **425**, 737–741.
32. Davis, I., Leaver-Fay, A., Chen, V., Block, J., Kapral, G., Wang, X. *et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383.
33. Laskowski, R., Macarthur, M., Moss, D. & Thornton, J. (1993). Procheck—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
34. Lima, C. D., D’Amico, K. L., Naday, I., Rosenbaum, G., Westbrook, E. M. & Hendrickson, W. A. (1997). MAD analysis of FHIT, a putative human tumor suppressor from the HIT protein family. *Structure*, **5**, 763–774.
35. Wedekind, J. E., Frey, P. A. & Rayment, I. (1996). The structure of nucleotidylated histidine-166 of galactose-1-phosphate uridylyltransferase provides insight into phosphoryl group transfer. *Biochemistry*, **35**, 11560–11569.
36. Otwinowski, Z. & Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. *Macromol. Crystallogr.* **276**, 307–326.
37. Brodersen, D. E., de La Fortelle, E., Vornrhein, C., Bricogne, G., Nyborg, J. & Kjeldgaard, M. (2000). Applications of single-wavelength anomalous dispersion at high and atomic resolution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **56**, 431–441.
38. Sheldrick, G. M. (2008). A short history of SHELX. *Acta Crystallogr., Sect. A*, **64**, 112–122.
39. Hao, Q., Gu, Y. X., Zheng, C. D. & Fan, H. F. (2000). OASIS: a computer program for breaking phase ambiguity in one-wavelength anomalous scattering or single isomorphous substitution (replacement) data. *J. Appl. Crystallogr.* **33**, 980–981.
40. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55**, 849–861.
41. Terwilliger, T. C. (2003). Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **59**, 38–44.
42. Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 1002–1011.
43. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N. *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 213–221.
44. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240–255.
45. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
46. Collaborative Computational Project, N (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 760–763.
47. DeLano, W. (2002). *The PyMOL Molecular Graphics System*. De Lano Scientific, San Carlos, CA.