Structure, Vol. 10, 589–600, April, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S0969-2126(02)00746-3

The Crystal Structure of Diadenosine Tetraphosphate Hydrolase from *Caenorhabditis elegans* in Free and Binary Complex Forms

Scott Bailey,[1] Svetlana E. Sedelnikova,[1] G. Michael Blackburn,[2] Hend M. Abdelghany,[3] Patrick J. Baker,[1] Alexander G. McLennan,[3] and John B. Rafferty[1,4]

[1]Krebs Institute for Biomolecular Research
Department of Molecular Biology and Biotechnology
University of Sheffield
Western Bank
Sheffield S10 2TN

[2]Krebs Institute for Biomolecular Research
Department of Chemistry
University of Sheffield
Western Bank
Sheffield S3 7HF

[3]School of Biological Sciences
University of Liverpool
Life Sciences Building
Liverpool L69 7ZB
United Kingdom

Summary

The crystal structure of *C. elegans* Ap_4A hydrolase has been determined for the free enzyme and a binary complex at 2.0 Å and 1.8 Å, respectively. Ap_4A hydrolase has a key role in regulating the intracellular Ap_4A levels and hence potentially the cellular response to metabolic stress and/or differentiation and apoptosis via the Ap_3A/Ap_4A ratio. The structures reveal that the enzyme has the mixed α/β fold of the Nudix family and also show how the enzyme binds and locates its substrate with respect to the catalytic machinery of the Nudix motif. These results suggest how the enzyme can catalyze the hydrolysis of a range of related dinucleoside tetraphosphate, but not triphosphate, compounds through precise orientation of key elements of the substrate.

Introduction

The Ap_4A hydrolase from the nematode *Caenorhabditis elegans* belongs to a family of enzymes that hydrolyze a diphosphate (pyrophosphate) linkage in dinucleoside 5′,5‴-*P*[1],*P*[n]-polyphosphates with four or more phosphate groups, such as diadenosine 5′,5‴-*P*[1],*P*[4]-tetraphosphate (Ap_4A), always producing an NTP as one of the products [1, 2]. Asymmetrically cleaving Ap_4A hydrolases (diadenosine 5′,5‴-*P*[1],*P*[4]-tetraphosphate pyrophosphohydrolases) have been isolated from a variety of animal and plant sources [2] and have recently been found in proteobacteria [3, 4], which were previously believed to possess only an unrelated symmetrically cleaving Ap_4A hydrolase that generates 2 moles of ADP from Ap_4A [5, 6].

[4]Correspondence: j.rafferty@sheffield.ac.uk

Almost all asymmetrically cleaving Ap_4A hydrolases characterized thus far are enzymes of the Nudix hydrolase family [1, 7]. Nudix hydrolases are a group of phosphoanhydrolases that predominantly catalyze hydrolysis of the diphosphate linkage in a variety of nucleoside triphosphates, dinucleoside polyphosphates, nucleotide sugars, and related compounds having the general structure of a *nu*cleoside *di*phosphate (NDP) linked to another moiety *X*, where X may be, for example, an NMP, NDP, P_i, or a sugar residue (where N is any nucleotide). Most members of the Nudix family possess the consensus sequence motif $GX_5EX_7REUXEEXGU$ (where U represents Ile, Leu, or Val, and X represents any amino acid), while others have a closely related signature sequence [7, 8], which forms part of the catalytic site for diphosphate hydrolysis. Interestingly, sequence alignments between animal and plant Ap_4A hydrolases reveal few clear similarities in regions outside the Nudix sequence motif itself [9], with the proteobacterial enzymes showing much greater similarity to the plant enzymes than to those from animals [3]. Furthermore, phylogenetic analysis of plant, animal, and prokaryotic enzymes confirms that plant and proteobacterial Ap_4A hydrolases form a group that is distinct from the animal enzymes [10]. Interestingly, this analysis also revealed that the enzymes from the archaea *Pyrobaculum aerophilum* and *Halobacterium halobium* that are predicted (but not experimentally confirmed) to be Ap_4A hydrolases are more similar to the animal enzymes than they are to the plant enzymes. Thus, it appears that even given their common catalytic properties, on the basis of primary sequence, the Ap_4A hydrolases can be divided into two distinct groups: the "animal-type" Ap_4A hydrolases that contain the subgroups of animal and archaeal sequences and the "plant-type" Ap_4A hydrolases that contain the subgroups of plant and proteobacterial sequences.

Ap_4A is synthesized predominantly by aminoacyl-tRNA synthetases by the addition of the AMP moiety from an aminoacyl-AMP to ATP [11]. The precise function of Ap_4A is still unclear. On the one hand, it may be an unavoidable by-product of protein synthesis that has to be cleared from the cell before it attains a potentially toxic concentration. In this case, Ap_4A hydrolase would perform the housecleaning role proposed to be one function of members of the Nudix hydrolase family [7]. On the other hand, Ap_4A has been implicated as a regulatory factor in DNA replication and as a component of the cellular response to metabolic stress, including oxidative stress and heat shock [12, 13].

Recent evidence further implicates Ap_4A in essential cellular functions, which would imply an equally important regulatory role for Ap_4A hydrolases [13, 14]. For example, several physiological and pathological effects have been found to be associated with a change in the intracellular Ap_3A/Ap_4A ratio. Differentiation and apopto-

Key words: Ap_4A hydrolase; asymmetric cleavage; magnesium cluster; Nudix family; substrate pocket; *Caenorhabditis elegans*

Table 1. Data Collection and Phasing Statistics

		Free Enzyme				Binary Complex	
Data set		Inflection	Peak	Remote	Native	Home Source	Synchrotron
Data collection statistics							
Wavelength Å		0.99988	1.00887	0.918407	0.87000	1.54179	0.87000
Resolution Å		18–2.05	18–2.05	18–2.05	18–1.98	20–2.01	20–1.8
(last shell)		(2.10–2.05)	(2.10–2.05)	(2.10–2.05)	(2.05–1.98)	(2.06–2.01)	(1.84–1.80)
Reflections (unique)		30,718 (8683)	31,483 (8925)	32,171 (8921)	53,439 (9864)	36,397 (16,982)	62,050 (24,019)
R_{merge} % (last shell)[a]		4.7 (8.9)	5.5 (10.4)	5.2 (10.5)	6.3 (23.8)	7.4 (27.3)	9.7 (33.9)
Completeness % (last shell)		99.6 (99.5)	99.6 (99.7)	99.6 (99.9)	97.9 (99.8)	86.3 (94.8)	97.0 (96.5)
I/sigma (last shell)		23.5 (9.2)	23.2 (8.7)	21.6 (7.7)	26.9 (9.5)	11.4 (3.1)	11.6 (3.7)
Phasing statistics							
Phasing power[b]	acentric	–	1.35	1.31			
	centric	–	1.07	1.00			
R_{cullis}[c]	acentric	–	0.74	0.75			
	centric	–	0.67	0.69			
	anomalous	0.61	0.61	0.58			
FoM[d]		0.58					

[a] $R_{merge} = \Sigma_{hkl}|I_i - I_m|/\Sigma_{hkl}I_m$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively.
[b] Phasing Power = <F_H/lack of closure>, where F_H is the calculated heavy-atom structure factor amplitude.
[c] R_{Cullis} = <lack of closure>/<isomorphous difference>
[d] FoM = mean figure of merit.

sis of cultured cells have significant and opposite correlations with this ratio, differentiation being associated with an increase in the ratio, and apoptosis with a decrease [15]. Furthermore, Ap_4A has been shown to induce apoptosis in several reversibly permeabilised human cell lines, whereas Ap_3A is a coinducer of the differentiation of HL60 cells [16]. Ap_4A may also be an important ligand of the Fhit tumor suppressor protein [14, 17]. In all the above schemes, Ap_4A hydrolase would play an important role in the regulation of Ap_4A levels and hence the Ap_3A/Ap_4A ratio.

A role for Ap_4A hydrolase has also been proposed in the invasive phenotype of pathogenic bacteria. A two-gene locus, *ialAB*, has been identified in the genome of *Bartonella bacilliformis*, the only bacterium known to invade erythrocytes. This locus confers an invasion phenotype on minimally invasive strains of *Escherichia coli* when introduced by transformation [18], and the *ialA* gene has recently been shown to encode a plant-type Ap_4A hydrolase, suggesting a role for Ap_4A metabolism in the invasion process [3, 4]. *Bartonella* spp. are now emerging as important human pathogens [19], and sequence homologs of *ialA* are present in the genomes of a number of other invasive bacteria [1]. If Ap_4A metabolism is required for optimal bacterial survival during invasion, then the product of *ialA*-related genes may be regarded as a new virulence factor. Attenuation of virulence is one option in the control of serious bacterial infection, so inhibitors that can discriminate between the bacterial plant-type Ap_4A hydrolase and the human animal-type homolog may be of value as new therapeutic agents.

The structures of three Nudix hydrolases have thus far been determined; they are *E. coli* MutT 8-oxo-dGTPase [20], *Lupinus angustifolius* Ap_4A hydrolase [21], and *E. coli* orf209 ADP-ribose pyrophosphatase [22]. These structures have revealed a characteristic Nudix fold [22] consisting of a mixed and an antiparallel β sheet and several α helices. The Nudix consensus sequence is located on a loop-helix-loop structural motif. Structures of ADP-ribose pyrophosphatase and MutT have also been determined in the presence of substrate [22] and competitive inhibitor [23], respectively. These structures reveal that, although the folds are very similar in the two proteins, both the mode of substrate binding and specific substrate contacts are different. Only the diphosphate moiety of the substrate is positioned equivalently. The rest of the substrate, most notably the base, binds on opposite sides of the catalytic site [22].

We have undertaken a structural investigation of the animal-type Ap_4A hydrolase from the nematode *C. elegans*. We present here the structure of the *C. elegans* enzyme from crystals grown in the presence and absence of the hydrolysis-resistant substrate analog, diadenosine 5′,5‴-(*P*^{2},*P*^{3}-methylene)-*P*^{1},*P*^{4}-tetraphosphate ($AppCH_2ppA$) [24]. The structure of an enzyme binary complex provides insights into the substrate binding and specificity of this group of Nudix hydrolases. A comparison of the structures with the NMR structure of the plant-type *L. angustifolius* enzyme reveals the differences between the two Ap_4A hydrolase subgroups. We have addressed the question of how, given their low sequence similarity outside the Nudix sequence motif, the two groups stabilize substrate and what is their possible evolutionary relationship. Furthermore, a comparison of the structure of the *C. elegans* Ap_4A hydrolase with the structures of the *E. coli* MutT and ADP-ribose pyrophosphatase has been performed that yields important insights into substrate binding and catalysis by the Nudix hydrolase family as a whole.

Results and Discussion

Structure Determination

Ap_4A hydrolase from *C. elegans* was overexpressed and purified as described elsewhere [25]. Crystals of the free enzyme belong to space group $C222_1$ [10] and have one monomer in the crystallographic asymmetric unit. The

Table 2. Refinement Statistics

	Free Enzyme	Binary Complex
Resolution (Å)	18–2	18–1.8
Final R factor (%)	20.1	18.9
R_{free} (%)	21.8	22.2
Rmsd bonds (Å)	0.012	0.009
Rmsd angles (°)	1.5	1.4
No. protein atoms in model	1013	2147
No. solvent atoms in model	132	320
Ramachandran analysis (%) (Most favored/additional allowed)	90.3/9.7	94.1/5.9
Average B values (Å^2)[a]		
Protein[b]	15	6
Water	34	19
AMP moiety	–	24
Magnesium ions	–	10
P^4 anion	–	10

[a] After TLS refinement.
[b] Average B factors for both chains A and B are the same for the binary complex

structure was solved by a multiwavelength anomalous dispersion (MAD) experiment (see Table 1 and Experimental Procedures) exploiting the anomalous signal of a mercurial derivative. The experimental map resulting from MAD phases was improved by density modification with solvent flattening and histogram matching. The final model containing 132 residues and 148 water molecules was refined against native data collected from a crystal at 2.0 Å resolution to an R factor of 20.1% and an R_{free} of 21.8%. Analysis of the geometry (Table 2) shows all parameters are inside the expected value range for a model at this resolution.

Crystals of the binary complex belong to space group $P2_1$ and have two monomers in the crystallographic asymmetric unit (Table 1). The structure was solved by molecular replacement (MR) using the refined free enzyme structure as a search model. Two molecules of the enzyme (termed chain A and chain B) were found. The initial MR phases were improved and bias removed by density modification with solvent flipping, density truncation, and two fold noncrystallographic symmetry averaging (see Experimental Procedures). The final model containing 137 residues of chain A, 132 residues of chain B, 322 water molecules, 9 magnesium ions, 2 anions, and an AMP moiety was refined to an R factor of 18.9% and R_{free} of 22.2%. Analysis of the geometry (Table 2) shows all parameters are inside the expected value range for a model at this resolution.

Structure Description

The *C. elegans* Ap_4A hydrolase molecule has dimensions approximately 40 × 35 × 30 Å with the overall architecture of the "Nudix fold" [22], consisting of a curved central four-stranded mixed β sheet (strand order βC′, βA, βE, and βD), a three-stranded antiparallel β sheet (strand order βF, βB, and βC), three α helices (αI–αIII), and the connecting loops (L1–L8) (Figure 1). The β sheets are sandwiched between helices αII and αIII (which form an antiparallel pair) and helix αI (Figure 1A). As in other Nudix hydrolases, the Nudix sequence motif (GX_5EX_7 REUXEEXGU; see above) is located in a loop-helix-loop motif formed by helix αI (the "catalytic helix") and its connecting loops, L3 and L4 (residues 37–59 in the *C. elegans* enzyme).

In the Ap_4A hydrolase binary complex crystals, the crystallographic asymmetric unit contains two molecules (chain A and chain B) that are related by a noncrystallographic two fold rotation. The two molecules are very similar, with root-mean-square (rms) deviations between 122 α carbon atomic positions of 0.7 Å. Gel filtration studies of *C. elegans* Ap_4A hydrolase suggest that the protein is monomeric [10].

The Substrate Binding Pocket

Examination of the electron density map for the binary complex of Ap_4A hydrolase with $AppCH_2ppA\text{-}Mg^{2+}$ revealed several density features in a surface cleft of chain A between the β sheets, previously associated with substrate binding to MutT and *L. angustifolius* Ap_4A hydrolase. Although there was no interpretable density for the entire $AppCH_2ppA$ molecule, the density for one of the adenosine groups and an associated phosphate was clearly identifiable (Figure 2A). This phosphate has been designated the P^1-phosphate in accordance with the

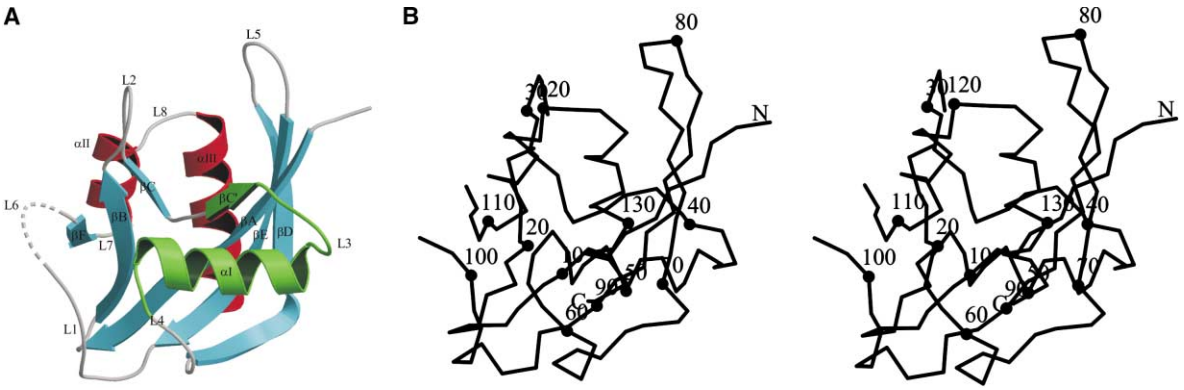

Figure 1. The Fold of *C. elegans* Ap4A Hydrolase

(A) A ribbon diagram showing the overall fold of the free enzyme with α helices and β strands shown as red coils and blue arrows, respectively, except for the region containing the Nudix motif, which is highlighted in green. The disordered L6 loop is shown dashed.
(B) A stereo image of the backbone of the enzyme with the position of every tenth residue labeled (produced using Molscript [49]).

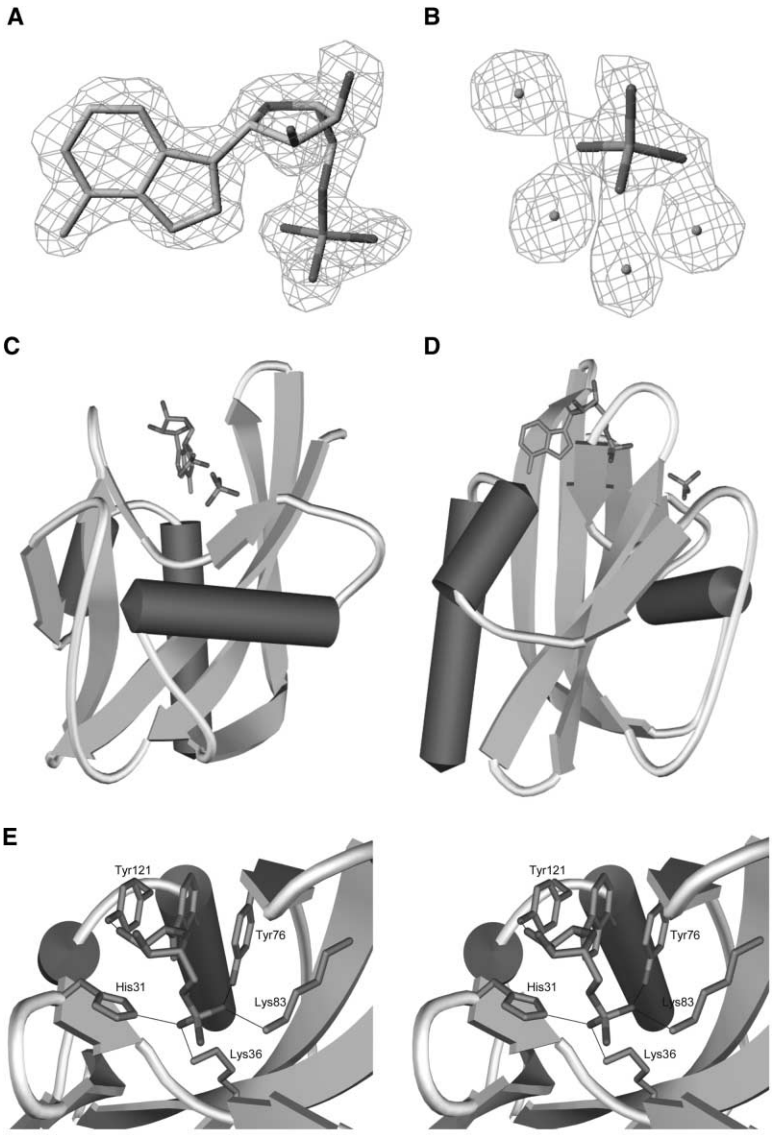

Figure 2. The Binary Complex

(A) The electron density map for the AMP moiety in the final 2Fo-Fc map contoured at 1σ.

(B) An Fo-Fc electron density omit map for the bound anion at the P^4-phosphate site and associated magnesium cations.

(C and D) Orthogonal views of the substrate binding cleft of the enzyme showing the location of the bound AMP moiety and the anion at the P^4 phosphate site (see text). The protein is shown with α helices and β strands as cylinders and arrows, respectively, and the AMP and P^4 anions are shown in all-atom representation.

(E) A stereo view of the binding site for the AMP moiety with the key interacting residues (see text) shown. Hydrogen bonds/ion pairs are shown as black lines. Figures produced using WebLabViewer V4.0 and TURBO-FRODO (A. Roussel et al., 1997, XV IUCr Congress, abstract).

convention proposed by Guranowski and coworkers [26] in which the P^1-phosphate is attached to the more strongly bound nucleotide. Additionally, the electron density in the region around helix αI (Figure 2B) that contains the Nudix sequence motif is very well defined and shows unambiguously the position of a single sulfate or phosphate anion, four magnesium ions, and a hydroxide ion (see Experimental Procedures for assignment details). The location of this anion relative to the catalytically essential glutamates of the Nudix sequence motif and its proximity to the P^1-phosphate have prompted its assignment as a probable position for the P^4-phosphate of the substrate Ap_4A that undergoes nucleophilic attack during catalysis. In the following discussion, it will be referred to as the P^4-phosphate [26, 27]. Further examination of the electron density remote from this region revealed an additional magnesium ion located on the interface formed between loop L7 and the N-terminal end of strand βB of chain A with the N-terminal end of a crystallographically related loop L6 of chain B. Inspection of chain B revealed equivalent features for the anion, magnesium, and hydroxide ions, but no clear density was observed for any other part of the $AppCH_2ppA$ molecule, including the AMP moiety found in chain A. Indeed, the density for the region around the substrate binding site, particularly in loops L5 and L8, of chain B is generally not as well defined as that of chain A. Therefore, the following discussion will be limited to observations within chain A.

The distance between the P^1- and P^4-phosphate binding sites identified in the electron density map is approximately large enough to accommodate both the P^2- and P^3-phosphates of the $AppCH_2ppA$ molecule, but too large for just one phosphate. The lack of interpretable electron density for the P^2- and P^3-phosphates could be accounted for in a number of ways. A possible scenario arises from the presence of intact $AppCH_2ppA$ in the crystal lattice, but such that there is sufficient flexibility in the location of the P^2- and P^3-phosphates and the second adenosine moiety to result in disorder and hence

no interpretable electron density. Given the well-defined electron density of the AMP moiety and the proposed P^4-phosphate, this seems unlikely. A more likely scenario is that the substrate analog has been turned over in the active site, as it is not totally refractory to hydrolysis by enzymes of this type [28, 29]. In this scenario, the observed AMP moiety is part of either the ATP product, where its P^2- and P^3-phosphates are disordered in the crystal lattice, or the AMP product itself, which has occupied the binding site after the ATP has diffused out. The anion at the P^4-phosphate site is then most likely to be a phosphate or sulfate ion carried over from the purification of the enzyme [25].

AMP Moiety Binding Pocket

Analysis of the AMP moiety in the binary complex reveals that the molecule sits in a groove on the protein surface formed by the two β sheets orientated such that the adenine moiety is directed toward helices αII and αIII and the phosphate toward the "catalytic helix" αI (Figures 2C and 2D). The nucleotide binds in a closed conformation (Figure 2A), with the ribose sugar found in the C2′-*endo* C3′-*exo* conformer and the adenine ring in the *anti* conformation.

The adenine ring binds in a pocket formed by the side chains of Tyr76 located at the C-terminal end of strand βD and Tyr121 situated in loop L8. The two phenolic rings of the tyrosines sandwich the adenine ring, forming extensive π-π stacking interactions, burying approximately 130 $Å^2$ (50%) of the solvent-accessible surface area of the adenine ring (Figure 2E). This double-stacking arrangement has also been observed in several other nucleotide binding proteins, including the unrelated mammalian cap binding translation initiation factor [30] and vaccinia virus mRNA-cap-dependent 2′-*O*-methyltransferase [31]. The N6 nitrogen of the adenine group extends into a pocket formed by the side chains of Thr33, Leu74, Val85, Tyr87, Ala119, and Met124, where it forms a direct hydrogen bond to the central of three bound water molecules. The size of the pocket and the lack of base-specific interactions are consistent with the ability of the enzyme to hydrolyze a variety of dinucleoside polyphosphate substrates [32]. The 2′ hydroxyl of the ribose moiety makes a hydrogen bond with the side chain hydroxyl group of Tyr121. In addition, the ribose oxygen O4 makes a van der Waals contact with the side chain phenolic ring of Tyr76 (Figure 2E). Stabilization of the P^1-phosphate moiety on the enzyme is achieved via a series of hydrogen bond/salt bridges between the phosphate oxygens and the side chain NZ nitrogen of Lys36 and Lys83, the side chain hydroxyl group of Tyr76, and the side chain imidazole ring Nϵ2 of His31 (Figure 2E).

P^4-Phosphate Binding Site

The proposed P^4-phosphate forms only one direct interaction with the protein via a hydrogen bond between one of its oxygens and the main chain amide of His38. However, it does contact several water molecules and magnesium ions. In total, four magnesium ions were found in the catalytic site above helix αI (Figure 3). Three of the magnesium ions are positioned so as to form an isosceles triangle (inter-ion distances 3.1 and 3.5 Å) at the base of the P^4-phosphate such that each one of three magnesium ions is ligated by one of the oxygens of the P^4-phosphate (Figure 3). The fourth magnesium ion lies out of the plane of the other three and 3.2 Å away from the nearest of them. Each magnesium is hexacoordinated, with ligands provided by the P^4-phosphate, the main chain carboxyl of Lys36, the side chain carboxyl groups of Glu52, Glu56, and Glu103 (from the Nudix sequence motif), a hydroxide ion, and nine other solvent molecules presumed to be water (Figure 3). Thus, the orientation of a bound P^4-phosphate is quite constrained through its interactions with the magnesium ions. The necessity for precise binding of a P^4-phosphate by the enzyme in order to correctly orientate its substrate for nucleophilic attack is consistent with the inability of Ap_4A hydrolase to catalyze the hydrolysis of an Ap_3A substrate [10, 32].

Studies on the metal requirements of the plant-type *B. bacilliformis* hydrolase [33] suggest a requirement for only two catalytic magnesium ions. Therefore, the observation of four magnesium ions (and one hydroxide ion) in the region of the "catalytic" helix in the *C. elegans* enzyme is unexpected. Crystals of the binary complex only grew in the presence of 200 mM $MgCl_2$, far in excess of the optimum 3–5 mM concentration for catalysis [10], and thus the additional ions may represent an artifact of crystallization. It is not possible to predict with confidence which of the four magnesium ions might be bound in an authentic substrate complex in vivo. However, the structure does support the role of residues Glu52, Glu56, and Glu103 as metal coordinating ligands, and mutation of these residues to glutamines results in 10^3-, 10^5-, and 40-fold reductions in K_{cat}, respectively (A.G.McL., unpublished data).

Second Nucleotide Binding Site

Examination of the region beyond the P^4-phosphate reveals, in the absence of significant conformational changes, no clear binding site for the second adenosine moiety. Ap_4A hydrolase will hydrolyze a range of Np_4X substrates [2, 32], and this is consistent with the lack of a defined binding site for the X moiety. This would suggest that the interactions of the protein with the Np_4 moiety of the substrate are sufficient to stabilize the substrate and orientate the P^4-phosphate for nucleophilic attack, as previously described [27].

P^2-P^3 Phosphate Binding Region

There was no interpretable electron density for either the P^2- or P^3-phosphates. However, inspection of the protein in this region did reveal that Lys83 is positioned such that it could stabilize the P^2- as well as the P^1-phosphate, and residues His38, Lys79, and Tyr27 might participate in phosphate stabilization either by direct interaction or via metal coordination. This latter suggestion is prompted by studies of the metal requirements of the plant-type *B. bacilliformis* Ap_4A hydrolase that have shown one further metal ion to bind the Ap_4A substrate in addition to the two catalytically required metal ions [33].

Substrate Analog-Induced Conformational Changes

Superposition of the structures of the free and binary complex forms of Ap_4A hydrolase shows that they have a very similar overall conformation. The value for the

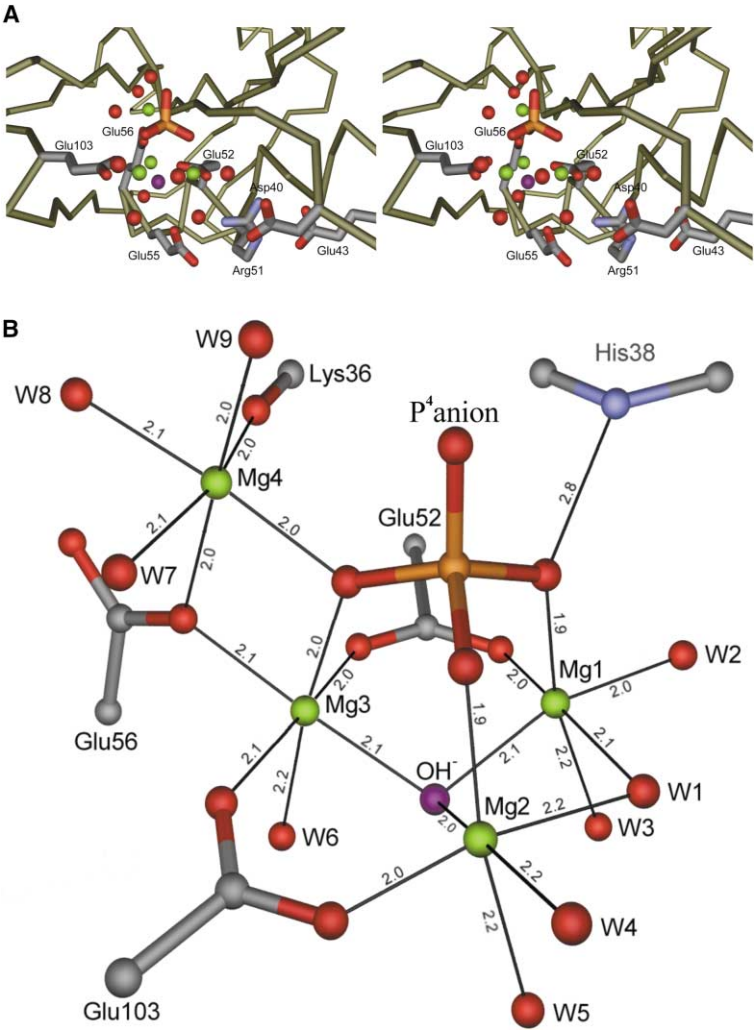

Figure 3. The Binding Site for the P^4-Phosphate

(A) A stereo view of the P^4-phosphate binding site above the "catalytic helix" αI showing the bound anion depicted as a phosphate in stick representation, associated Mg^{2+} cations (green spheres), hydroxide ion (purple sphere), water molecules (red spheres), and coordinating protein side chain ligand network (see text). The protein backbone is shown in dark green.

(B) Detailed view of the anion coordination scheme showing the novel magnesium cluster with atoms colored as in (A) and ligand bonds shown as black lines with distances in angstroms marked. (Produced using WebLabViewer v4.0.)

rmsd of 101 out of 138 α carbon positions (residues 2–24, 32–75, 82–94, and 106–136) is 0.4 Å. The largest changes in main chain positions occur in three loop regions (Figure 4A). First, the loops L2 (residues 25–31) and L5 (residues 74–82) that flank the adenosine binding site undergo small conformational changes in order to accommodate substrate binding. Second, the disordered loop, L6 (residues 93–106), which had no interpretable density in the free enzyme structure, becomes ordered upon substrate and metal binding, with well-defined density unambiguously outlining its path. The ordered L6 loop is stabilized via the coordination of the Oϵ1 and Oϵ2 atoms of Glu103 to two of the four magnesium ions located in the catalytic site. Closer examination of the side chain positions within the substrate binding site reveals several significant alterations (Figure 4B). A rotation of approximately 90° about the χ1 dihedral angle of Tyr121 occurs with an associated shift in the side chain of Tyr76 such that the planes of the two phenolic rings lie parallel and at a distance of approximately 7 Å and results in the formation of the binding pocket for an adenine ring of the substrate. Movement of the Tyr121 side chain is therefore essential for localization of an adenine ring on the enzyme, although the trigger for the conformational change has yet to be determined. The side chain of His31 rotates approximately 90° about its χ1 dihedral angle. This reorientation results in the side chain imidazole ring Nϵ2 atom coming into hydrogen bonding distance of the O1 oxygen of the P^1-phosphate. Examination of the residues of the Nudix sequence motif reveals that, with the exception of Glu56, none of the residues shows a significant shift. Glu56, however, does shift the position of its side chain carboxyl group to optimize its coordination geometry to two of the four magnesium ions located in the catalytic site. The only other significant shifts in side chain positions occur in regions where the crystal contacts differ between the two crystal forms.

Comparison with *L. angustifolius* Ap_4A Hydrolase

Superposition of the *L. angustifolius* plant-type Ap_4A hydrolase with the structures of both the free and substrate analog bound *C. elegans* animal-type Ap_4A hydrolase both gave rmsds of 1.6 Å over 91 matched α carbon positions (Figure 5A). Analysis of this superposition reveals the two folds are closely equivalent. The major differences in the structures are a 7 residue insertion in loop L6 and an extra, ill-defined helix (helix αII)

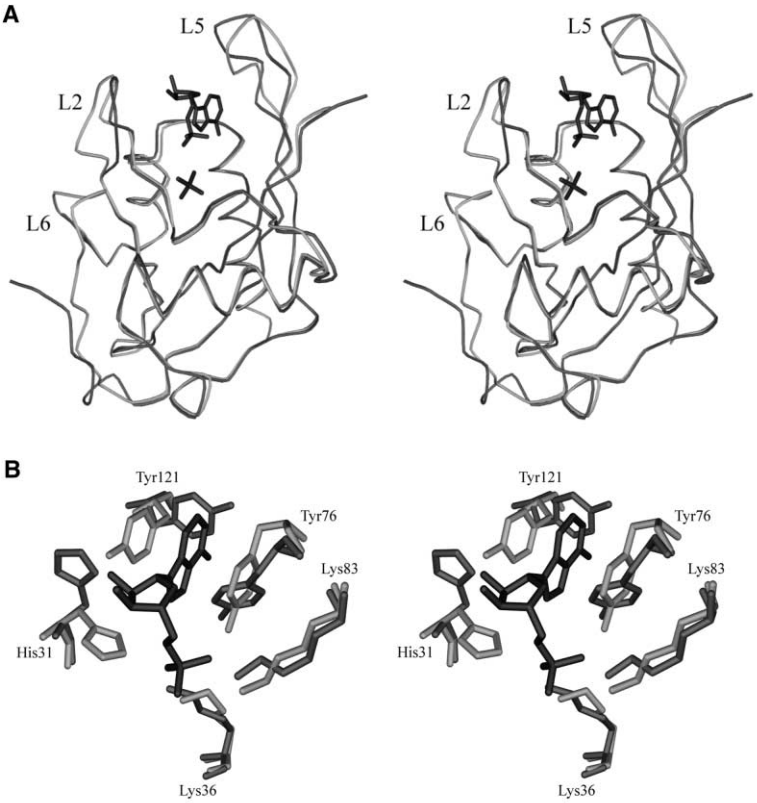

Figure 4. Comparison of the Fold of the Enzyme in the Free and Binary Complex Forms

(A) A stereo view of the superimposition of the backbone trace of the free enzyme (dark gray) with that of the enzyme in the binary complex (light gray) showing the conformational changes involved in substrate binding. The bound AMP moiety and anion are shown for reference (black stick representation).

(B) A stereo view of the details of the conformational changes at the binding site for the AMP moiety showing the relevant affected protein side chains and the AMP moiety. Shading as in (A). (Produced using WebLabViewer V4.0.)

in loop L5 of the *L. angustifolius* enzyme [21]. In the plant enzyme, loop L2 adopts a conformation that results in a much more open substrate binding pocket. Loop L2 in *C. elegans* forms one side of the binding pocket and supplies residue His31 for interaction with the P^1-phosphate of the substrate. Additionally, loop L6 in the *L. angustifolius* enzyme, which contains Glu125, the only catalytically essential glutamate not within the Nudix sequence motif, is found in a conformation reminiscent of the *C. elegans* binary complex. Thus in the *L. angustifolius* enzyme, the catalytically essential residue Glu125 [34] is positioned toward the catalytic site even in the absence of substrate, whereas the equivalent residue in *C. elegans*, Glu103, is localized to the catalytic site only upon substrate binding (see above).

Alignment of the plant-type and animal-type sequences based on the structural superposition of the *C. elegans* and *L. angustifolius* structures results in the sequence alignment shown in Figure 5B. Although there is conservation within the "animal" or "plant" groups, inspection of the alignment reveals that, outside the Nudix sequence motif, only 5 residues (Trp32, Pro35, Tyr76, Glu103, and Trp109 in *C. elegans*) are completely conserved across the whole Ap_4A hydrolase family. Of these, Tyr76 is involved in a stacking interaction with one of the substrate adenine rings, and its role could be replicated by Tyr77 in the *L. angustifolius* enzyme. The Tyr121 from loop L8 that forms the other face of the adenine binding pocket is not completely conserved. Its role might be replicated most obviously by the use of Phe144 in the *L. angustifolius* enzyme, as suggested by both sequence and structural alignments, although the conformation adopted by loop L8 is notably different in the two structures and a proper substrate binding analysis is required. Among the other completely conserved residues, Glu103 ligates the bound metal, but none of the other three residues forms interactions with the substrate in the *C. elegans* binary complex structure. The tryptophan residues are required presumably to stabilize the protein fold, and Pro35 helps orientate the C-terminal end of strand βC, which contains the Nudix sequence motif residue Gly37, parallel to the catalytic helix αI. Indeed, not only does there seem to be a general lack of conservation in the substrate binding pocket, but also there is a marked difference in charge distribution within the substrate binding sites of the *C. elegans* and *L. angustifolius* enzymes (Figure 5C) that might perhaps reflect differences in magnesium cation binding. In addition, there are noticeably more polar residues in the binding site of the *L. angustifolius* enzyme (Figure 5C).

It has been reported that a number of residues in the *L. angustifolius* enzyme are sensitive to titration with a substrate analog [21]. These residues were mapped onto the superposition of the *C. elegans* and *L. angustifolius* enzymes to give further insights into the mode in which each enzyme binds substrate. Perhaps the most interesting observation is that several residues (Asn88, Gly92, Ser93, Trp91, and Trp95 in *L. angustifolius*) sensitive to the titration are located around the C-terminal end of the αII helix of the plant enzyme, which has no equivalent in the *C. elegans* enzyme and might suggest a role in substrate binding for this region unique to the plant-type Ap_4A hydrolases.

The existence of two distinct groups of Ap_4A hy-

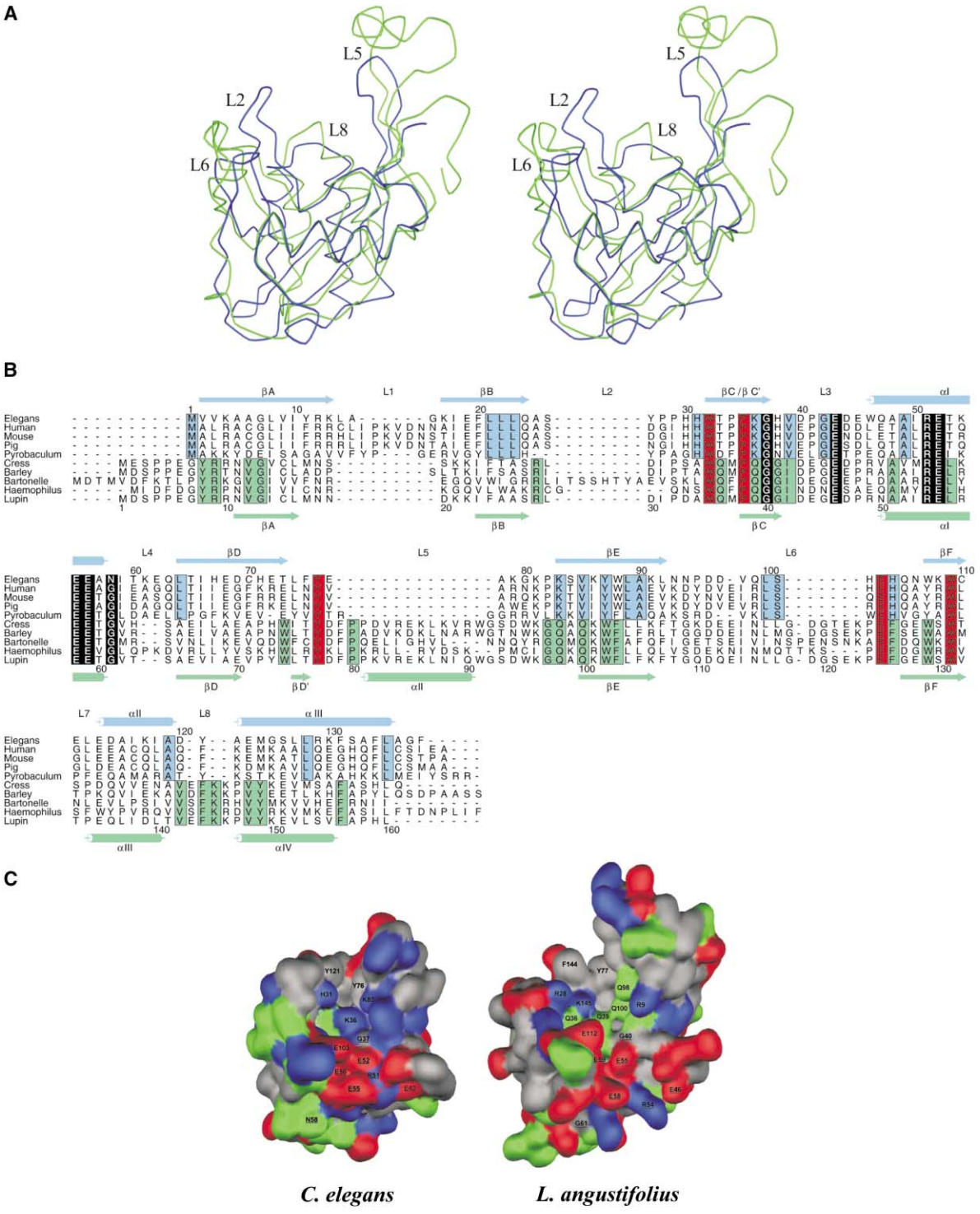

Figure 5. A Comparison between the *C. elegans* and *L. angustifolius* Ap_4A Hydrolases and Other Family Members

(A) A stereo view of the superimposition of the backbone trace of the binary complex form of the *C. elegans* enzyme (blue) with that of the *L. angustifolius* enzyme (green).

(B) A structure-based sequence alignment of the members of the two groups, animal-type and plant-type, of Ap_4A hydrolase. Residues have been highlighted on the basis of the Nudix consensus sequence (reverse type face), conservation across all species outside the Nudix motif (red box), and conservation among animal-type (blue box) or plant-type (green box) sequences.

(C) Surface representation of the *C. elegans* and *L. angustifolius* Ap_4A hydrolases showing the difference in shape and distribution of charge within the substrate binding site. Positively charged (including His), negatively charged, and polar residues are colored blue, red, and green, respectively, while the rest are gray. The conserved residues within each group (animal-type or plant-type) as well as the Nudix motif residues (underlined) are labeled. (Figures produced using WebLabViewer V4.0.)

drolases raises the question of their evolution. Whether the lack of conservation in the substrate binding pocket and elsewhere in the sequences outside the Nudix motif is a reflection of divergence of an ancestral Ap_4A hydrolase or the convergence of separate Nudix hydrolases toward the same substrate, Ap_4A, remains unclear.

Comparison with Other Nudix Hydrolases

There are structures available for two other members of the Nudix hydrolase family, the NMR structure of the *E. coli* nucleoside triphosphate pyrophosphohydrolase, MutT, with and without a nonhydrolysable substrate analog [20, 23], and the crystal structure of *E. coli orf209* ADP-ribose pyrophosphatase (ADPRase), with and without ADP-ribose [22]. Both of these enzymes catalyze the hydrolysis of the P^1-P^2 linkage of their respective substrates. (In this discussion, the α, β, and γ phosphates of the substrates described in the literature for MutT and ADPRase have been redefined as P^1, P^2, and P^3 for consistency with the Ap_4A hydrolase description.) They both have a very similar overall fold to the *C. elegans* Ap_4A hydrolase, although the dimeric nature of ADPRase results in a binding pocket that contains residues from both subunits and a noticeably different orientation for the substrate.

In the structures of the *C. elegans* Ap_4A hydrolase and MutT, the bound adenosine moiety of their substrate analogs lies against the equivalent β strands (βA, βD, and βE in Ap_4A hydrolase and βA, βC, and βD in MutT), although in MutT it is exposed to solvent on the other face, whereas in Ap_4A hydrolase the adenine ring is buried in a pocket. The relative positions of the P^2-phosphate in MutT and putative P^4-phosphate of Ap_4A hydrolase are very similar, lying close to the catalytic helix αI and coordinated by magnesium ions bound by residues from the Nudix sequence motif. In contrast, the bound ADP-ribose substrate in the ADPRase complex resembles a "horseshoe" [22] and is thus orientated such that the direction of its P^1-P^2 phosphate linkage is opposite to that of the substrate analog in the MutT complex. Nonetheless, the diphosphate links that undergo hydrolysis in all three enzymes are positioned very similarly in space relative to the catalytic helix αI and the residues of the Nudix sequence motif. Thus, the basic structural architecture provided by the common mixed α/β fold of these enzymes has been modified so as to present the appropriate diphosphate bonds in the correct positions to the catalytic machinery of the Nudix residues for hydrolysis.

Catalytic Mechanism of *C. elegans* Ap_4A Hydrolase

The proposed mechanism of hydrolysis of the diphosphate link catalyzed by Ap_4A hydrolase requires the activation of a bound solvent molecule by a base to facilitate its nucleophilic attack on the P^4-phosphate of the substrate [26, 27]. Examination of the region around the putative P^4-phosphate in the binary complex, the mutation data on the *C. elegans* and *L. angustifolius* Ap_4A hydrolases, and data on *E. coli* MutT (A.G.McL., unpublished data) [28, 35] suggests Glu56 as the most likely candidate for the catalytic base. Work on the *L. angustifolius* Ap_4A hydrolase has shown an inversion of stereochemistry upon hydrolysis [36], consistent with an *exo* pattern of attack by the activated water [26]. However, in the absence of an intact substrate or analog and the presence of the unexpectedly large number of bound magnesium ions, it is not possible to predict with confidence which, if any, of the bound solvent molecules might act as the attacking nucleophile.

Biological Implications

The enzyme Ap_4A hydrolase catalyzes the breakdown of Ap_4A, a ubiquitous molecule produced mainly as a by-product of protein synthesis. The asymmetric cleavage of Ap_4A carried out by the *C. elegans* enzyme studied here results in the production of ATP and AMP, which can be recycled into the cellular stored energy pool. Thus, the enzyme may have a simple housekeeping function, but the Ap_4A molecule has been implicated in playing a role in one or more essential cellular processes, including DNA replication, metabolic stress, and apoptosis, and thus the hydrolase may have a more important function in the careful regulation of the intracellular levels of Ap_4A. In particular, the ratio of Ap_4A to Ap_3A has been linked to the choice between differentiation and apoptosis in cultured cell lines. The *C. elegans* Ap_4A hydrolase belongs to the Nudix family and is the first monomeric member to have its crystal structure determined. The structure of the enzyme both free and in a binary complex shows how a conformational change accompanies substrate binding and how one nucleoside and the P^1- and P^4-phosphates are bound. The conformational change brings key conserved residues into position to enable catalysis to occur. The structure reveals how the enzyme can discriminate between dinucleoside tetraphosphate and triphosphate substrates and hence can help specifically in maintaining the Ap_4A/Ap_3A ratio. Furthermore, the lack of any direct, base-specific contacts to the bound nucleoside and the absence of an obvious second nucleoside binding site offer an explanation for the ability of the enzyme to act upon a range of nucleoside tetraphosphate-containing substrates. Structure and sequence comparisons show how the *C. elegans* enzyme and other eukaryotic and archaeal Ap_4A hydrolases form a subclass with structural characteristics distinct from those of the plant and proteobacterial enzymes.

Experimental Procedures

Protein Crystallization

C. elegans Ap_4A hydrolase was expressed and purified as described elsewhere [25]. Crystals were grown in the presence and absence of $AppCH_2ppA$-Mg^{2+} by the hanging drop method of vapor diffusion. The free enzyme crystals belong to space group $C222_1$ with cell dimensions a = 62.8 Å, b = 72.9 Å, and c = 61.1 Å and were grown with well solutions containing 32% (w/v) PEG4000, 0.2 M ammonium acetate and 100 mM sodium citrate (pH 5.6). V_M calculations suggested a solvent content of 43% for a monomer in the asymmetric unit. Heavy-atom derivatives of these crystals were prepared by soaking the crystals in 2 mM ethyl mercury phosphate (EMP) for 2 hr. The Ap_4A hydrolase binary complex was produced by addition of 10 mM $AppCH_2ppA$ and 10 mM magnesium chloride to the *C. elegans* enzyme. The crystals were grown by equilibration against

well solutions of 30% (w/v) PEG4000, 0.2 M magnesium chloride, and Tris-HCl (pH 8.5) and belong to space group $P2_1$ with cell dimensions a = 57.6 Å, b = 36.8 Å, and c = 68.9 Å with a β = 114.2°. V_M calculations suggested a solvent content of 40% for two monomers in the asymmetric unit. For data collection, both crystal forms were cooled straight from the drop in a stream of nitrogen gas at 100 K.

Data Collection and Processing

A full three-wavelength MAD data experiment was performed on a single EMP-soaked free enzyme crystal, and data were collected to a minimum Bragg spacing of 2.0 Å using a Mar Research CCD scanner at the ESRF station BM14. Three wavelengths were chosen near the Hg absorption edge (Table 1) based on an X-ray fluorescence scan of the frozen crystal in order to maximize the f′′ component (λ_1, peak), to minimize the f′ component (λ_2, inflexion), and to maximize Δf′ (λ_3, remote). The data for each wavelength were processed individually and scaled in such a way as to preserve anomalous signal using the HKL Suite of programs [37]. Native data for the free enzyme were collected from a single frozen crystal on an ADSC Quantum 4 CCD camera at the Daresbury SRS station 9.6 to 2.0 Å resolution. Details of the data collection statistics are given in Table 1.

Data for the binary complex were collected from a single frozen crystal both on a Mar345 image plate scanner mounted on a Rigaku RU200 X-ray generator with Yale focusing mirror optics to 2.0 Å resolution and subsequently on an ADSC Quantum 4 CCD camera at the Daresbury SRS station 9.6 to 1.8 Å resolution. Data for each dataset were processed using the HKL Suite of programs. Details of the data collection statistics are given in Table 1.

Phasing, Model Building, and Structure Refinement of the Free Enzyme Structure

The positions of four mercury atoms were found in the asymmetric unit of the EMP-soaked free enzyme data using the direct methods procedure, as implemented in the program SnB [38]. Data for the peak wavelength were scaled and processed with the DREAR program package [39] to derive the renormalized E values corresponding to the anomalous differences. These E values were then submitted to SnB using the program's standard parameters. Analysis of the minimum function values showed a bimodal distribution indicating convergence to a solution.

The mercury atom parameters were then refined, and initial phases were calculated in the program MLPHARE [40] following the pseudo-MIR procedure [41]. The phases were further improved by solvent flattening and histogram matching using structure factor amplitudes from the λ_3 (remote) dataset with the program DM [42]. The resulting 2.0 Å map was of excellent quality, with readily identifiable side chain density, and therefore these modified phases were used in the automated model building procedure implemented by the program ARP/wARP [43]. This procedure produced a polyalanine chain trace corresponding to 75% of the main chain, and this initial model then formed the basis of a manual build in the program TURBO-FRODO (A. Roussel et al., 1997, XV IUCr Congress, abstract). The resulting model was subjected to restrained maximum-likelihood refinement at a resolution of 2.0 Å, in the program REFMAC [44], against data collected from a native crystal. Iterative cycles of phase combination of the partial structure phases and those from the MAD experiment, model building, and refinement, which in the later stages was performed using TLS parameterisation to account for the overall anisotropic motion of the protein [45], were used to construct a model representing 132 residues of the 138 expected residues, with 92% of these residues assigned unambiguously. ARP/wARP also added a total of 109 solvent molecules automatically during the refinement process, and tight control was maintained over the geometric parameters of the model, including the main chain torsion angles as assessed by Ramachandran plot analysis using the program PROCHECK [46]. There was no interpretable electron density for residues Ser26, Glu104, Gln105, and the final two C-terminal residues, Gly137 and Phe138.

Molecular Replacement of the Binary Complex

The structure of the binary complex was solved by molecular replacement (MR) with the program AMoRe [47], using the refined structure of the free enzyme as a search model. The two subunits in the asymmetric unit were successfully located from a solution with a correlation of 58.5 and an R factor of 43.7%. Analysis of this MR solution in TURBO-FRODO revealed consistent crystal packing. A $3F_o$-$2F_c$ electron density map was calculated on the MR solution using data from 20–2 Å resolution that indicated a correct solution had been found but showed poor electron density around several loop regions. Therefore, the calculated phases were subjected to density modification within the program CNS [48], which included solvent flipping, density truncation, and two fold NCS averaging. The resulting density-modified map showed significant improvements, with loop regions now clearly defined. The model was rebuilt to fit the observed density and then submitted to restrained maximum-likelihood refinement using REFMAC. Iterative cycles of model building and refinement were used to construct a model representing 137 and 132 of the 138 expected residues (with 93% of these residues assigned unambiguously) in chain A and chain B, respectively. At this point, density within the expected substrate binding sites was accounted for in chains A and B by the addition of a phosphate ion (a sulfate ion can also be equally well refined at this position), four magnesium ions (assigned on the basis of their hexacoordination with excellent octahedral geometry and short ligand bonds of average length 2.1 Å), and a hydroxide ion (assigned as the most appropriate species present able to coordinate three magnesium ions simultaneously with bond lengths of approximately 2.1 Å) as well as the addition of an AMP molecule to the binding-site of chain A. The model was then further refined, and a total of 320 solvent molecules were added by cycles of REFMAC and ARP/wARP. At this time, the higher resolution data became available (1.8 Å), and subsequent refinement was done using this data. In the final stages, one additional magnesium ion was added remote from either substrate binding site, and TLS parameterisation was used to account for the overall anisotropic motion of each protein chain in the asymmetric unit [45]. Tight control was maintained of the geometric parameters of the model including the main chain torsion angles, as assessed by Ramachandran plot analysis using the program PROCHECK [46]. No restraints were used for the magnesium-ligand distances. The refined temperature factors of the bound ligands were similar to those of the surrounding protein molecule prior to TLS parameterization, and thus all the ligand molecules are considered as being fully occupied in both chains A and B. There was no interpretable electron density for residues Met1 of chain A and the first three N-terminal residues (Met1, Val2, and Val3), Thr9, Gly80, and Lys81 of chain B.

Acknowledgments

We are grateful for the support of Hassan Belrhali and the station scientists on beamline BM14 at the ESRF and station 9.6 at the CLRC Daresbury SRS laboratory and for travel funds provided by EMBL under the EC Human Capital and Mobility Program. H.M.A. is the recipient of a postgraduate scholarship from the Egyptian government. Work in the laboratories of J.B.R. and A.G.McL. is supported by the BBSRC and The Wellcome Trust. J.B.R. is a Royal Society Olga Kennard Fellow.

Received: October 24, 2001
Revised: January 9, 2002
Accepted: January 23, 2002

References

1. McLennan, A.G. (1999). The MutT motif family of nucleotide phosphohydrolases in man and human pathogens (review). Int. J. Mol. Med. *4*, 79–89.
2. Guranowski, A. (2000). Specific and nonspecific enzymes involved in the catabolism of mononucleoside and dinucleoside polyphosphates. Pharmacol. Ther. *87*, 117–139.
3. Cartwright, J.L., Britton, P., Minnick, M.F., and McLennan, A.G. (1999). The *lal*A invasion gene of *Bartonella bacilliformis* encodes a (di)nucleotide polyphosphate hydrolase of the MutT motif family and has homologs in other invasion bacteria. Biochem. Biophys. Res. Commun. *256*, 474–479.

4. Conyers, G.B., and Bessman, M.J. (1999). The gene, *ial*A, associated with the invasion of human erythrocytes by *Bartonella bacilliformis*, designates a nudix hydrolase activity on dinucleoside 5′-polyphosphates. J. Biol. Chem. *274*, 1203–2354.
5. Guranowski, A., Jakubowski, H., and Holler, E. (1983). Catabolism of diadenosine 5′,5‴-P^1,P^4-tetraphosphate in prokaryotes. Purification and properties of diadenosine 5′,5‴-P^1,P^4-tetraphosphate (symmetrical) pyrophosphohydrolase from *Escherichia coli* K12. J. Biol. Chem. *258*, 14784–14789.
6. Plateau, P., Fromant, M., Brevet, A., Gesquière, A., and Blanquet, S. (1985). Catabolism of bis(5′-nucleosidyl) oligophosphates in *Escherichia coli*: metal requirements and substrate specificity of homogeneous diadenosine 5′,5‴-P1,P4-tetraphosphate pyrophosphohydrolase. Biochemistry *24*, 914–922.
7. Bessman, M.J., Frick, D.N., and O'Handley, S.F. (1996). The MutT proteins or 'nudix' hydrolases, a family of versatile, widely distributed, "housecleaning" enzymes. J. Biol. Chem. *271*, 25059–25062.
8. Koonin, E.V. (1993). A highly conserved sequence motif defining the family of MutT-related proteins from eubacteria, eukaryotes and viruses. Nucleic Acids Res. *21*, 4847.
9. Maksel, D., Guranowski, A., Ilgoutz, S.C., Moir, A., Blackburn, M.G., and Gayler, K.G. (1998). Cloning and expression of diadenosine 5′,5‴-P^1,P^4-tetraphosphate hydrolase from *Lupinus angustifolius* L. Biochem. J. *329*, 313–319.
10. Abdelghany, H.D., Gasmi, L., Cartwright, J.L., Bailey, S., Rafferty, J.B., and McLennan, A.G. (2001). Cloning, characterization and crystallization of a diadenosine 5′-5‴-P^1,P^4-tetraphosphate pyrophosphohydrolase from *Caenorhabditis elegans*. Biochim. Biophys. Acta, in press.
11. Goerlich, O., Foeckler, R., and Holler, E. (1982). Mechanism of synthesis of adenosine(5′)tetraphospho(5′)adenosine (AppppA) by aminoacyl-tRNA synthetases. Eur. J. Biochem. *126*, 135–142.
12. Baxi, M.D., and Vishwanatha, J.K. (1995). Diadenosine polyphosphates: their biological and pharmacological significance. J. Pharmacol. Toxicol. Methods *33*, 121–128.
13. Kisselev, L.L., Justesen, J., Wolfson, A.D., and Frolova, L.Y. (1998). Diadenosine oligophosphates (Ap_nA), a novel class of signaling molecules? FEBS Lett. *427*, 157–163.
14. McLennan, A.G. (2000). Dinucleoside polyphosphates - friend or foe? Pharmacol. Ther. *87*, 73–89.
15. Vartanian, A., Prudovsky, I., Suzuki, H., DalPra, I., and Kisselev, L. (1997). Opposite effects of cell differentiation and apoptosis on Ap_3A/Ap_4A ratio in human cell cultures. FEBS Lett. *415*, 160–162.
16. Vartanian, A., Alexandrov, I., Prudowski, I., McLennan, A., and Kisselev, L. (1999). Ap_4A induces apoptosis in cultured human cells. FEBS Lett. *456*, 175–180.
17. Murphy, G.A., Halliday, D., and McLennan, A.G. (2000). The Fhit tumor suppressor protein regulates the intracellular concentration of diadenosine triphosphate but not diadenosine tetraphosphate. Cancer Res. *60*, 2342–2344.
18. Mitchell, S.J., and 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.
19. Anderson, B.E., and Neumann, M.A. (1997). *Bartonella* spp. as emerging human pathogens. Clin. Microbiol. Rev. *10*, 203–219.
20. Abeygunawardana, C., Weber, D.J., Gittis, A.G., Frick, D.N., Lin, J., Miller, A.F., Bessman, M.J., and Mildvan, A.S. (1995). Solution structure of the MutT enzyme, a nucleoside triphosphate pyrophosphohydrolase. Biochemistry *34*, 14997–15005.
21. Swarbrick, J.D., Bashtannyk, T., Maksel, D., Zhang, X., Blackburn, G.M., Gayler, K.R., and Gooley, P.R. (2000). The three-dimensional structure of the Nudix enzyme diadenosine tetraphosphate hydrolase from *Lupinus angustifolius* L. J. Mol. Biol. *302*, 1165–1177.
22. Gabelli, S.B., Bianchet, M.A., Bessman, M.J., and Amzel, L.M. (2001). The structure of ADP-ribose pyrophosphate reveals the structural basis for the versatility of the Nudix family. Nat. Struct. Biol. *8*, 467–472.
23. Lin, J., Chitrananda, J.L., Frick, D.N., Bessman, M.J., and Mildvan, A.S. (1997). Solution structure of the quaternary MutT-M^{2+}-AMPCPP-M^{2+} complex and mechanism of its pyrophosphohydrolase action. Biochemistry *36*, 1199–1210.
24. Blackburn, G.M., Guo, M.-J., and McLennan, A.G. (1992). Synthetic structural analogues of dinucleoside polyphosphates. In Ap_4A and Other Dinucleoside Polyphosphates, A.G. McLennan, ed. (Boca Raton, Florida: CRC Press), pp. 305–342.
25. Bailey, S., Sedelnikova, S.E., Blackburn, G.M., Abdelghany, H.M., McLennan, A.G., and Rafferty, J.B. (2002). Crystallization of a complex of *Caenorhabditis elegans* diadenosine tetraphosphate hydrolase and a non-hydrolysable substrate-analogue, $AppCH_2ppA$. Acta Crystallogr. *D58*, 526–528.
26. Guranowski, A., Brown, P., Ashton, P.A., and Blackburn, G.M. (1994). Regiospecificity of the hydrolysis of diadenosine polyphosphates catalyzed by three specific pyrophosphohydrolases. Biochemistry *33*, 235–240.
27. McLennan, A.G., Prescott, M., and Evershed, R.P. (1989). Identification of point of specific enzymic cleavage of P^1,P^4-bis(5′-adenosyl) tetraphosphate by negative ion FAB mass spectrometry. Biomed. Environ. Mass Spectrom. *18*, 450–452.
28. Guranowski, A., Starzynska, E., Taylor, G.E., and Blackburn, G.M. (1989). Studies on some specific Ap_4A-degrading enzymes with the use of various methylene analogues of P^1,P^4-bis-(5′,5‴-adenosyl) tetraphosphate. Biochem. J. *262*, 241–244.
29. McLennan, A.G., Taylor, G.E., Prescott, M., and Blackburn, G.M. (1989). Recognition of $\beta\beta'$-substituted and $\alpha\beta$, $\alpha'\beta'$-disubstituted phosphonate analogues of bis(5′-adenosyl) tetraphosphate by the bis(5′-nucleosidyl)-tetraphosphate pyrophosphohydrolases from *Artemia* embryos and *Escherichia coli*. Biochemistry *28*, 3868–3875.
30. Marcotrigiano, J., Gingras, A.C., Sonnenberg, N., and Burley, S.K. (1997). Cocrystal structure of the messenger RNA 5′ cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell *89*, 951–961.
31. Hodel, A.E., Gershon, P.D., and Quiocho, F.A. (1998). Structural basis for sequence nonspecific recognition of 5′-capped mRNA by a cap modifying enzyme. Mol. Cell *1*, 443–447.
32. Guranowski, A., and Sillero, A. (1992). Enzymes cleaving dinucleoside polyphosphates. In Ap_4A and Other Dinucleoside Polyphosphates, A.G. McLennan, ed. (Boca Raton, Florida: CRC Press), pp. 81–133.
33. Conyers, G.B., Wu, G., Bessman, M.J., and Mildvan, A.S. (2000). Metal requirements of a diadenosine pyrophosphatase from *Bartonella bacilliformis*: magnetic resonance and kinetic studies of the role of Mn^{2+}. Biochemistry *39*, 2347–2354.
34. Maksel, D., Gooley, P.R., Swarbrick, J.D., Guranowski, A., Gange, C., Blackburn, G.M., and Gayler, K.R. (2001). Characterization of active-site residues in diadenosine tetraphosphate hydrolase from *Lupinus angustifolius*. Biochem. J. *357*, 399–405.
35. Lin, J., Abeygunawardana, C., Frick, D.N., Bessman, M.J., and Mildvan, A.S. (1996). The role of Glu57 in the mechanism of the *Escherichia coli* MutT enzyme by mutagenesis and heteronuclear NMR. Biochemistry *35*, 6715–6726.
36. Dixon, R.M., and Lowe, G. (1989). Synthesis of (R_p,R_p)-P^1,P^4-Bis(5′-adenosyl)-1[^{17}O,$^{18}O_2$]tetraphosphate from (S_p,S_p)-P^1,P^4-Bis(5′-adenosyl)-1[*thio*-$^{18}O_2$], 4[*thio*-$^{18}O_2$]tetraphosphate with retention at phosphorus and the stereochemical course of hydrolysis by the unsymmetrical Ap_4A phosphodiesterase from lupin seeds. J. Biol. Chem. *264*, 2069–2074.
37. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode (San Diego: Academic Press).
38. Weeks, C.M., DeTitta, G.T., Hauptman, H.A., Thuman, P., and Miller, R. (1994). Structure solution by minimal-function phase refinement and Fourier filtering. II. Implementation and applications. Acta Crystallogr. A *50*, 210–220.
39. Blessing, R.H., Guo, D.Y., and Langs, D.A. (1998). Intensity statistics and normalization. In Direct Methods for Solving Macromolecular Structures, NATO ASI Series Volume, Series C: Mathematica and Physical Sciences, Volume 507, S. Fortier, ed. (Dordrecht, The Netherlands: Kluwer Academic Publishers), pp. 47–71.
40. Otwinowski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. In Proceedings of the CCP4 Study Weekend,

W. Wolf, P.R. Evans, and A.G.W. Leslie, eds. (Warrington, UK: SERC Daresbury Laboratory), pp. 80–88.
41. Ramakrishnan, V., and Biou, V. (1997). Treatment of multiwavelength anomalous diffraction data as a special case of multiple isomorphous replacement. In Methods in Enzymology: Macromolecular Crystallography, Volume 276, C.W. Carter, and R.M. Sweet, eds. (San Diego: Academic Press), pp. 538–557.
42. Cowtan, K. (1994). "DM": an automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography *31*, pp. 34–38.
43. Lamzin, V.S., and Wilson, K.S. (1993). Automated refinement of protein models. Acta Crystallogr. D *49*, 129–149.
44. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D *53*, 240–255.
45. Winn, M.D., Isupov, M., and Murshudov, G.N. (2001). Use of TLS parameters to model anisotropic displacement in macromolecular refinement. Acta Crystallogr. D *57*, 122–133.
46. Laskowski, R.A., Macarthur, M.W., Moss, D.W., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.
47. Navazza, J. (1994). AmoRe: an automated package for molecular replacement. Acta Crystallogr. A *47*, 392–400.
48. Brunger, A.T., and Warren, G.L. (1998). Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. Acta Crystallogr D *54*, 905–921.
49. Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Accession Numbers

Coordinates of the free enzyme and binary complex have been deposited in the RCSB with accession codes 1KT9 and 1KTG.