Published in final edited form as:
J Mol Biol. 2014 January 9; 426(1): . doi:10.1016/j.jmb.2013.09.025.

Crystal structure of human poly(A) polymerase gamma reveals a conserved catalytic core for canonical poly(A) polymerases

Qin Yang[a], **Lydia Nausch**[a], **Georges Martin**[b], **Walter Keller**[b], and **Sylvie Doublié**[a,*]

[a]Department of Microbiology and Molecular Genetics, University of Vermont, Stafford Hall, 95 Carrigan Drive, Burlington VT 05405-0068, USA [b]Computational and Systems Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Abstract

In eukaryotes, the poly(A) tail added at the 3′ end of a mRNA precursor is essential for the regulation of mRNA stability and the initiation of translation. Poly(A) polymerase (PAP) is the enzyme that catalyzes the poly(A) addition reaction. Multiple isoforms of PAP have been identified in vertebrates, which originate from gene duplication, alternative splicing, or post-translational modifications. The complexity of PAP isoforms suggests that they might play different roles in the cell. Phylogenetic studies indicate that vertebrate PAPs are grouped into three clades termed α, β and γ, which originated from two gene duplication events. To date, all the available PAP structures are from the PAPα clade. Here, we present the crystal structure of the first representative of the PAPγ clade, human PAPγ, bound to cordycepin triphosphate (3′dATP) and Ca^{2+}. The structure revealed that PAPγ closely resembles its PAPα ortholog. An analysis of residue conservation reveals a conserved catalytic binding pocket, whereas residues at the surface of the polymerase are more divergent.

Keywords

mRNA processing; 3′end processing; polyadenylation; poly(A) polymerase gamma; neo-PAP; cordycepin triphosphate

Introduction

3′ end processing of transcripts generated by RNA polymerase II is a fundamental event for the maturation of messenger RNA in eukaryotes [1; 2]. To date, twenty 3′end processing factor subunits have been identified in yeast [3; 4], and more than 80 proteins have been co-purified with 3′end processing factors in human cells, although the exact number of factors directly involved in processing is currently unknown [5]. The complexity of the 3′ processing machinery ensures a seamless crosstalk with other steps of gene expression [6; 7] and precise regulation of 3′ processing [8; 9; 10; 11; 12; 13]. The majority of eukaryotic pre-mRNAs 3′ ends

*Corresponding author: sdoublie@uvm.edu, Phone: +1-802-656-9531, Fax: +1-802-656-8749.

Accession number
Coordinates and structure factors have been deposited in the Protein Data Bank with accession number 4LT6.

are processed in two intimately coupled steps: an initial site-specific endonucleolytic cleavage followed by the addition of a poly(A) tail at the 3′ end of the upstream cleavage product [4; 14]. The poly(A) tail has crucial functional importance for enhancing mRNA stability and translational regulation. The length of poly(A) tail appears to be a key determinant in deciding the fate of polyadenylated RNAs [15; 16; 17]. Poly(A) polymerase (PAP) is the enzyme that synthesizes the poly(A) tail [18; 19; 20], but it also stimulates the cleavage reaction through interactions with other 3′ end processing factors [14; 21; 22].

PAP is a template-independent RNA polymerase that belongs to the X family of polymerases, which is characterized by the signature helix-turn motif hG[G/S]$x_{9\text{-}13}$Dh[D/E]h (x, any amino acid; h, hydrophobic amino acid) [23]. Crystal structures of bovine PAPα and yeast Pap1 revealed that PAP has three globular domains: a central domain flanked by an N-terminal catalytic domain, where the signature helix-turn motif is located, and a C-terminal domain, which shares similar topology with the ubiquitous RNA recognition motif (RRM) [24; 25; 26; 27]. A recent crystal structure of yeast Pap1 with oligo(A) and ATP revealed that the nucleotide and the last three nucleotides of the mRNA primer are bound within the substrate-binding cleft formed by the three domains [28]. Vertebrate PAPs contain an additional C-terminal domain (CTD), which is missing in yeast Pap1 [18; 19]. The presence of dual nuclear localization signals, multiple cyclin-dependent kinase phosphorylation sites and a splicing factor U1A interaction motif in the CTD are consistent with the regulatory role of PAP in gene expression (Figs. 1A and S1) [4; 23; 29].

Despite the monocatalytic function of PAP, multiple isoforms of this enzyme are observed in cell lines and tissues of several species ([30] and references within). In human cells, three genes have been identified, *PAPOLA* on chromosome 14 [31], an intronless *PAPOLB* on chromosome 7 [29], and *PAPOLG* on chromosome 2 [32; 33; 34], which encode the PAP isoforms, PAPα, PAPβ and PAPγ (also called neo-PAP), respectively. Human PAPα and PAPγ are nuclear enzymes [34], whereas the testis-specific PAPβ is found in the cytoplasm [29]. Furthermore, these PAPs are subject to alternative splicing and post-translational modifications ([30] and references within). The functional significance of the presence of multiple PAPs *in vivo* is not well understood, but it has been proposed that this multiplicity of enzyme forms ensures precise control of polyadenylation via interactions with other 3′ processing factors in different tissues and/or cell growth states [32; 35].

With the rapid development of sequencing techniques and the availability of genome information for a variety of species, we sought to investigate PAP genes in other vertebrates to further our understanding of the biological significance of organisms harboring multiple PAPs. In this report, our phylogenetic study illustrates that PAPs are grouped into three clades, α, β and γ, which originated via two gene duplication events. In light of the fact that all available PAP structures were from the PAPα clade, we set out to solve the crystal structure of human PAPγ (hPAPγ) bound to an ATP analog and divalent cation. The structure of hPAPγ is very similar to that of bovine PAPα (bPAPα), with the most highly conserved residues located in the active site cavity. Sequence analyses indicate that the C-terminal domain of the enzyme is more divergent and predicted to be intrinsically disordered.

Results and discussion

Sequence comparison of human PAPα, PAPβ, and PAPγ

A sequence alignment of three full-length human PAP isoforms and yeast Pap1 was carried out with T-Coffee [36]. The sequence conservation for the catalytic domain, the central domain and the RNA recognition motif, which collectively will be referred to as the N-terminal domain (NTD), is much higher than for the C-terminal domain (CTD) (Figs. 1A

and S1). hPAPβ and hPAPγ NTDs are 92% and 77% identical to hPAPα, whereas the CTDs are 63% and 28% identical. This observation supports the hypothesis that multiple PAPs execute their catalytic function through the conserved NTD, but are regulated via their distinct CTDs [32; 33; 34]. Indeed, while hPAPγ and hPAPα display undistinguishable polyadenylation activities, they appear to be differentially phosphorylated which suggests that they may be differentially regulated [32; 33; 34]. Compared to the dual nuclear localization signals (NLS) located in the CTD of hPAPα, only one NLS is present in hPAPβ [26], while three are found in hPAPγ [31]. The copy number of the NLS correlates well with the previous observation that hPAPγ is exclusively located in the nucleus [34], whereas hPAPβ is found primarily in the cytoplasm [29]. Furthermore, the U1A interaction motif and consensus phosphorylation sites for cyclin dependent kinases (SPXK/R) are conserved between hPAPα and hPAPγ, but are missing in hPAPβ. The biological significance of this observation needs further investigation. Interestingly, yeast Pap1 lacks a CTD but harbors an N-terminal extension of about 18 amino acids not seen in mammalian PAPs. This extension was recently demonstrated to play a part in the regulation of poly(A) addition through its interaction with other RNA processing factors [37] and might therefore function similarly to the CTD in higher eukaryotes.

PAPs are grouped into three clades, α, β, and γ

A phylogenetic analysis was carried out to gain a global view of the complexity of PAP. The vertebrate PAP phylogeny shows a gene duplication (Fig. 1B, circle), following the divergence of vertebrates from arthropods and nematodes but prior to the divergence of fish and tetrapods, that resulted in the PAPγ and PAPα+β clades. A second gene duplication (Fig. 1B, square), following the divergence of birds and mammals but preceding the divergence of rodents and primates, gave rise to the mammalian PAPα and PAPβ clades. Only one PAP sequence was found in the completely sequenced pufferfish (*Fugu rubripes)* genome [38]. In contrast, mammalian species have multiple clades of PAPs, which illustrates their biological importance.

Human PAPγ in complex with 3′dATP and Ca^{2+}

Multiple crystal structures are available for the PAPα clade [24; 25; 26; 28; 39]. In order to gain a structure-function understanding of the various PAP clades, we set out to determine the crystal structure of a member of the PAPγ clade. Initial attempts to crystallize full-length human PAPγ (residues 1-736) failed, probably due to the disordered nature of its CTD. A catalytically active CTD truncation construct (hPAPγ508, residues 1-508) [34] was designed based on a disorder prediction plot, which predicted residues 500 and higher to be disordered (Figure S2; [40]). hPAPγ508 was crystallized in complex with 3′deoxyadenosine 5′triphosphate, cordycepin triphosphate (3′dATP), a chain-terminating analog of ATP, and Ca^{2+} as the divalent cation. The crystals belong to space group $P2_12_12_1$, with cell dimensions a = 68.98 Å, b = 89.99 Å, and c = 202.08 Å. The structure of the hPAPγ508-3′dATP complex was solved by molecular replacement with the 2.15 Å model of bPAPα (PDB ID: 1Q79) [24] as the search model and refined to 2.8 Å with good statistics (Table S1). Two molecules are present in the asymmetric unit. One 3′dATP and one Ca^{2+} are present in the active site of each PAP molecule (Fig. 2).

The structure of hPAPγ508 exhibits the same tripartite architecture as bovine PAPα and yeast Pap1: an N-terminal catalytic domain, a central domain, and a C-terminal RNA binding domain with a fold reminiscent of an RRM [24; 25; 26; 28]. The ATP analog is bound in the cleft at the junction of the catalytic and central domains (Fig. 2A) and oriented in a conformation that would allow nucleophilic attack, as observed earlier in a Mg-ATP bPAPα complex (PDB ID: 1Q78 [24]) (Fig. 2B). Furthermore, the residues contacting the 3′dATP are identical to those described in bPAPα (Fig. 2B) [24]. Two strictly conserved catalytic

aspartates, Asp112 and Asp114, interact with Ca^{2+}, which also ligates three non-bridging oxygens of the α, β and γ phosphates of 3′dATP.

The triphosphate tail is further stabilized by the highly conserved residues Ser101, Tyr236, and Lys227 via hydrogen bonding and salt bridge interactions. The ribose moiety is sandwiched between Phe99 and Val246. The 2′OH is bound by a water molecule, which also interacts with Thr206. The adenine base stacks against Val246. The N1 position of the base is contacted by Thr316 (Fig. 2B) and C2 is within van der Waals distance of Val205. This valine is positioned to hinder GTP binding, as the extracyclic amino group of the purine would likely clash with the valine side chain.

Two Mn^{2+} ions were observed in bPAPα (PDB ID: 1Q79 [24]) and yPap1 (PDB ID: 1FA0 [26]), corresponding to metals A and B in DNA polymerase β, the founding member of the X-family of polymerases. The nucleotide-binding metal B coordinates the triphosphate tail of the incipient nucleotide whereas catalytic metal A contacts the 3′-OH of the primer, thereby making it a better nucleophile [41]. Only one Ca^{2+} was observed in the hPAPγ structure, corresponding to metal B, as was reported for bPAPα (PDB ID: 1Q78 [24]) and yPap1 (PDB ID: 2Q66 [28]) in complex with Mg^{2+}. We fully expect PAPγ to use the two-metal ion mechanism employed by DNA and RNA polymerases [42]. We may see only one metal because of the medium resolution of the crystallographic data (2.8Å), the presence of Ca^{2+} in the crystallization solution instead of Mg^{2+}, the likely physiological metal, or the absence of a free 3′OH, which is known to destabilize binding of metal ion A [41].

PAPs possess a conserved catalytic cavity and a divergent surface

The similitude in the overall shape of the NTD and the similarity of the residues participating in substrate binding between PAPα and PAPγ are consistent with the undistinguishable biochemical function of these two enzymes *in vitro* [32; 34]. Interestingly, by plotting the conservation of PAP residues onto the NTD domain, we observed a distinct pattern between the residues located in the catalytic cavity versus those at the surface of the protein (Fig 3A). The residues located in the substrate-binding pocket are highly conserved, whereas the residues at the surface of the enzyme are more divergent. A sequence alignment with yeast Pap1 reveals that most of the RNA-binding residues observed in the yPap1-RNA structure are mostly conserved in PAPγ (Figure S1). The surface representation in Figure 3B shows that the RNA binds in a crevice lined by highly conserved residues.

Crystallographic evidence has shown that the less conserved surface area of the C-terminal RNA binding domain of yeast Pap1 interacts with at least one of the other 3′ processing factors, Fip1 [21] (Fig. 3B). The less conserved region of the mammalian PAP CTD could similarly accommodate different protein-protein interaction patterns in the various PAP forms. This prediction is in agreement with the fact that most of the predicted post-translational modification sites are found, or predicted to be, in the CTD of PAPs α and γ [43; 44].

Comparison with other PAP structures

Rigid body domain movement is a component of the induced-fit catalytic mechanism proposed for PAP [28; 39; 45]. As previously reported for DNA polymerases, PAPs undergo a conformation change leading to domain closure and assembly of active site residues upon binding of the RNA primer and correct incoming nucleotide, Mg-ATP [46]. It should be noted that the conformational change is less pronounced than that reported for replicative DNA polymerases [42; 47]. hPAPγ was superimposed with bPAPα based on the central domain (residues 17-58 & 173-351)(Fig. 4)[39] to illustrate the movements of the catalytic and RRM domains. The rms deviation between the two models is 1.198 Å (comparing Cαs in PDB ID

4LT6 (this study) and 1Q78 [24]), and 1.465 Å comparing 4LT6 and 1Q79 [24; 31], indicating that the structures are overall quite similar but that there are regions that differ between the two models. A closer inspection reveals that PAPγ is in a slightly more closed conformation as compared to the bPAPα structures (Fig. 4A). The catalytic domain moves perpendicular whereas the RRM moves roughly parallel to the substrate binding cleft, as previously reported for yPap1 (Fig. 4B)[39]. Upon domain closure, the nucleotide triphosphate tail is contacted by three conserved residues, Ser101, Lys227, and Tyr236 (Fig. 2B). The latter two residues correspond to Lys215 and Tyr224 in yPap1, whose interactions with the triphosphate moiety are only observed in the closed state [28]. One difference with the yeast enzyme is that Asn238 in hPAPγ is located 6 Å from the adenine base, compared to ~3 Å for the analogous Asn226 in the nucleotide-bound closed form of yPap1 (Similarly Asn239 in bPAPα lies far from the nucleotide, about 8 Å). In the ternary complex of yPap1 with oligo(A) and MgATP Asn226 hydrogen bonds with the phosphate of the RNA primer terminal base and is within van der Waals distance of the adenine base [28]. hPAPγ would be predicted to further close in the presence of both nucleotide and oligo(A). A detailed analysis of the interaction of the conserved aspargine with the RNA primer and nucleotide will have to await a structure of the human enzyme with both substrates.

Conclusions

We present here the first crystal structure of PAPγ, obtained in complex with 3′dATP and Ca^{2+}. The polymerase closes around the nucleotide and makes conserved interactions to the triphosphate tail. Our phylogenetic and structural analyses indicate that all three clades of canonical PAP use the same conserved catalytic core for the polyadenylation reaction, whereas the less conserved CTDs are more likely to contribute to the diverse cellular function of PAPs and the regulation of 3′end processing.

Materials and methods

Sequence alignment and phylogenetic analysis

Sequences containing the poly(A) polymerase central domain (Conserved Domain Database [48] entry pfam04928.5) and the poly(A) polymerase RRM (pfam04926.5) were identified by CDART [49] and BLAST [50] searches of genome databases. Sequences were aligned with T-Coffee [36] and manually edited by SeaView [51] to remove poorly aligned regions. Preliminary phylogenetic reconstruction with Markov Chain Monte Carlo (MCMC) simulation as implemented in MrBayes [52] showed strong support for a clade of vertebrate sequences, represented in Figure 1B.

The sub-alignment corresponding to the vertebrate clade was manually filtered to remove redundant sequences and poorly aligned regions. The resulting alignment included 774 aligned sequence positions. Phylogenies were constructed by MCMC simulation using MrBayes (burnin=100000, ngen=510000, aamodelpr=fixed(jones). Phylogenies were also constructed by maximum likelihood estimation as implemented in PHYLIP [53]. Bootstrap support was based on 100 replicates

Protein expression and purification

A hPAPγ cDNA was isolated from reverse transcribed poly(A) containing RNA from Hela cells. The coding region of hPAPγ was PCR amplified with specific primers and the resulting sequence was confirmed to correspond to the reference sequence NM_022894. The PCR product was cloned into a plasmid vector and residues 1-508 were subcloned into expression vector pGM10 with a His_6 tag at the N-terminus [54; 55].

The C-terminally truncated human PAP γ (residues 1-508; hPAPγ508) carrying an N-terminal His tag was expressed in *E. coli* BL21 (DE3) cells (Novagen). Newly transformed cells were grown at 37°C in rich medium containing 50 μg ml^{-1} ampicillin (RPI Corporation). Expression was induced at an A_{600} of 0.5 by adding IPTG to a final concentration of 1 mM. Cells were grown overnight at room temperature. After lysis, hPAPγ508 was purified by Ni-NTA Superflow (Qiagen) chromatography followed by anion exchange chromatography (Resource Q column; Amersham Biosciences). A salt gradient ranging from 100 mM to 1 M KCl (20 mM Tris-HCl pH 8.0, 10% (v/v) glycerol, 4 mM βME) was applied. The peak fractions corresponding to hPAPγ508 were collected and concentrated to 5–10 mg ml^{-1} in 30K Amicon Ultra Centrifugal Devices (Millipore), flash frozen in liquid N_2 and stored at −80°C.

Crystallization and data collection

Crystals of hPAPγ508 in complex with 3′dATP were produced by mixing concentrated protein (5–7 mg ml^{-1}) containing 0.8 mM 3′dATP (in 10 mM HEPES pH 7) with crystallization buffer (20% (w/v) PEG 8000, 50 mM Tris-HCl pH 8.5, 100 mM $(NH_4)_2SO_4$, and 5 mM $CaCl_2$) in a 1:1 ratio. The resulting drop was equilibrated against a well of 1 ml crystallization solution at 18°C by the sitting-drop vapor-diffusion method. Crystals grew within two days to a size of about 0.15 × 0.04 × 0.04 mm^3 and belong to space group $P2_12_12_1$. The cell dimensions are a = 68.98 Å, b = 89.99 Å, c = 202.08 Å. There are two molecules per asymmetric unit, with an estimated solvent content of 57%. Crystals were cryoprotected overnight by addition of 5% (v/v) PEG 400 and by raising the PEG 8000 concentration to 13% (w/v) in the drop. Crystals were flash cooled in liquid nitrogen. Data were collected at the University of Vermont at 100 K on a Mar345 image plate detector (MarResearch) mounted on a rotating anode RU-300 X-ray generator (Rigaku) equipped with Xenocs mirrors. Data were indexed and processed with Denzo and Scalepack [56].

Structure determination and refinement

A molecular replacement solution with the previously determined 2.15 Å model of bovine PAP (PDB ID: 1Q79) [24] devoid of all non-protein atoms was found with Molrep [57]. Model building and further interpretation of the electron density map were performed with the program COOT [58]. Structural refinement was completed with Phenix.refine [59]. The current model consists of 471 residues for molecule A (17-422, 431-443, 454-501) and of 469 residues for molecule B (18-422, 431-443, 454-497). A molecule of 3′dATP and one calcium ion are found in the active site of both structures. The sequence conservation was calculated by the ConSurf server [60] with pre-aligned sequences from the phylogenetic analysis. All molecular structure figures were generated with PyMOL [15].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors gratefully acknowledge Drs. Ramiro Barrantes-Reynolds and Jeffrey Bond for help with the phylogenetic study, and Dr. Mark Rould for help with crystallographic data analysis. This work was supported by NIH award GM62239 to S. D. W. K. and G. M. were supported by the University of Basel, the Swiss National Science Fund (grants Nr. 3100-053897 and 31003A-143977) and the Louis-Jeantet Foundation for Medicine.

References

1. Mandel CR, Bai Y, Tong L. Protein factors in pre-mRNA 3′-end processing. Cell Mol Life Sci. 2008; 65:1099–122. [PubMed: 18158581]

2. Yang Q, Doublié S. Structural biology of poly(A) site definition. Wiley Interdiscip Rev RNA. 2011; 2:732–47. [PubMed: 21823232]
3. Proudfoot N, O'Sullivan J. Polyadenylation: a tail of two complexes. Curr Biol. 2002; 12:R855–7. [PubMed: 12498707]
4. Zhao J, Hyman L, Moore C. Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol Rev. 1999; 63:405–45. [PubMed: 10357856]
5. Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates JR 3rd, Frank J, Manley JL. Molecular architecture of the human pre-mRNA 3′ processing complex. Mol Cell. 2009; 33:365–76. [PubMed: 19217410]
6. Moore MJ, Proudfoot NJ. Pre-mRNA processing reaches back to transcription and ahead to translation. Cell. 2009; 136:688–700. [PubMed: 19239889]
7. Elkon R, Ugalde AP, Agami R. Alternative cleavage and polyadenylation: extent, regulation and function. Nat Rev Genet. 2013; 14:496–506. [PubMed: 23774734]
8. Millevoi S, Vagner S. Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation. Nucleic Acids Res. 2010; 38:2757–74. [PubMed: 20044349]
9. Yang Q, Coseno M, Gilmartin GM, Doublié S. Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. Structure. 2011; 19:368–77. [PubMed: 21295486]
10. Yang Q, Gilmartin GM, Doublié S. Structural basis of UGUA recognition by the Nudix protein CFIm25 and implications for a regulatory role in mRNA 3′ processing. Proc Natl Acad Sci U S A. 2010; 107:10062–10067. [PubMed: 20479262]
11. Yang Q, Gilmartin GM, Doublié S. The structure of human cleavage factor I(m) hints at functions beyond UGUA-specific RNA binding: a role in alternative polyadenylation and a potential link to 5′ capping and splicing. RNA Biol. 2011; 8:748–53. [PubMed: 21881408]
12. Tian B, Manley JL. Alternative cleavage and polyadenylation: the long and short of it. Trends Biochem Sci. 2013; 38:312–20. [PubMed: 23632313]
13. Franceschetti M, Bueno E, Wilson RA, Tucker SL, Gomez-Mena C, Calder G, Sesma A. Fungal Virulence and Development Is Regulated by Alternative Pre-mRNA 3 ′ End Processing in Magnaporthe oryzae. Plos Pathogens. 2011; 7
14. Wahle E, Ruegsegger U. 3′-End processing of pre-mRNA in eukaryotes. FEMS Microbiol Rev. 1999; 23:277–95. [PubMed: 10371034]
15. Kuhn U, Gundel M, Knoth A, Kerwitz Y, Rudel S, Wahle E. Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. J Biol Chem. 2009; 284:22803–14. [PubMed: 19509282]
16. Meyer S, Temme C, Wahle E. Messenger RNA turnover in eukaryotes: pathways and enzymes. Crit Rev Biochem Mol Biol. 2004; 39:197–216. [PubMed: 15596551]
17. Kuhn U, Wahle E. Structure and function of poly(A) binding proteins. Biochim Biophys Acta. 2004; 1678:67–84. [PubMed: 15157733]
18. Lingner J, Kellermann J, Keller W. Cloning and expression of the essential gene for poly(A) polymerase from S. cerevisiae. Nature. 1991; 354:496–8. [PubMed: 1840648]
19. Raabe T, Bollum FJ, Manley JL. Primary structure and expression of bovine poly(A) polymerase. Nature. 1991; 353:229–34. [PubMed: 1896071]
20. Wahle E, Martin G, Schiltz E, Keller W. Isolation and expression of cDNA clones encoding mammalian poly(A) polymerase. EMBO J. 1991; 10:4251–7. [PubMed: 1756732]
21. Meinke G, Ezeokonkwo C, Balbo P, Stafford W, Moore C, Bohm A. Structure of yeast poly(A) polymerase in complex with a peptide from Fip1, an intrinsically disordered protein. Biochemistry. 2008; 47:6859–69. [PubMed: 18537269]
22. Pezet-Valdez M, Fernandez-Retana J, Ospina-Villa JD, Ramirez-Moreno ME, Orozco E, Charcas-Lopez S, Soto-Sanchez J, Mendoza-Hernandez G, Lopez-Casamicha M, Lopez-Camarillo C, Marchat LA. The 25 kDa Subunit of Cleavage Factor Im Is a RNA-Binding Protein That Interacts with the Poly(A) Polymerase in Entamoeba histolytica. PLoS One. 2013; 8:e67977. [PubMed: 23840799]

23. Martin G, Doublie S, Keller W. Determinants of substrate specificity in RNA-dependent nucleotidyl transferases. Biochim Biophys Acta. 2008; 1779:206–16. [PubMed: 18177750]
24. Martin G, Moglich A, Keller W, Doublié S. Biochemical and structural insights into substrate binding and catalytic mechanism of mammalian poly(A) polymerase. J Mol Biol. 2004; 341:911–25. [PubMed: 15328606]
25. Martin G, Keller W, Doublié S. Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. EMBO J. 2000; 19:4193–203. [PubMed: 10944102]
26. Bard J, Zhelkovsky AM, Helmling S, Earnest TN, Moore CL, Bohm A. Structure of yeast poly(A) polymerase alone and in complex with 3′-dATP. Science. 2000; 289:1346–9. [PubMed: 10958780]
27. Clery A, Sinha R, Anczukow O, Corrionero A, Moursy A, Daubner GM, Valcarcel J, Krainer AR, Allain FHT. Isolated pseudo-RNA-recognition motifs of SR proteins can regulate splicing using a noncanonical mode of RNA recognition. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110:E2802–E2811. [PubMed: 23836656]
28. Balbo PB, Bohm A. Mechanism of poly(A) polymerase: structure of the enzyme-MgATP-RNA ternary complex and kinetic analysis. Structure. 2007; 15:1117–31. [PubMed: 17850751]
29. Lee YJ, Lee Y, Chung JH. An intronless gene encoding a poly(A) polymerase is specifically expressed in testis. FEBS Lett. 2000; 487:287–92. [PubMed: 11150526]
30. Colgan DF, Murthy KG, Prives C, Manley JL. Cell-cycle related regulation of poly(A) polymerase by phosphorylation. Nature. 1996; 384:282–5. [PubMed: 8918882]
31. Tupler R, Perini G, Green MR. Expressing the human genome. Nature. 2001; 409:832–3. [PubMed: 11237001]
32. Topalian SL, Kaneko S, Gonzales MI, Bond GL, Ward Y, Manley JL. Identification and functional characterization of neo-poly(A) polymerase, an RNA processing enzyme overexpressed in human tumors. Mol Cell Biol. 2001; 21:5614–23. [PubMed: 11463842]
33. Perumal K, Sinha K, Henning D, Reddy R. Purification, characterization, and cloning of the cDNA of human signal recognition particle RNA 3′-adenylating enzyme. J Biol Chem. 2001; 276:21791–6. [PubMed: 11287430]
34. Kyriakopoulou CB, Nordvarg H, Virtanen A. A novel nuclear human poly(A) polymerase (PAP), PAP gamma. J Biol Chem. 2001; 276:33504–11. [PubMed: 11431479]
35. Juge F, Zaessinger S, Temme C, Wahle E, Simonelig M. Control of poly(A) polymerase level is essential to cytoplasmic polyadenylation and early development in Drosophila. EMBO J. 2002; 21:6603–13. [PubMed: 12456666]
36. Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol. 2000; 302:205–17. [PubMed: 10964570]
37. Ezeokonkwo C, Ghazy MA, Zhelkovsky A, Yeh PC, Moore C. Novel interactions at the essential N-terminus of poly(A) polymerase that could regulate poly(A) addition in Saccharomyces cerevisiae. FEBS Lett. 2012; 586:1173–8. [PubMed: 22575652]
38. Kai W, Kikuchi K, Tohari S, Chew AK, Tay A, Fujiwara A, Hosoya S, Suetake H, Naruse K, Brenner S, Suzuki Y, Venkatesh B. Integration of the genetic map and genome assembly of fugu facilitates insights into distinct features of genome evolution in teleosts and mammals. Genome Biol Evol. 2011; 3:424–42. [PubMed: 21551351]
39. Balbo PB, Toth J, Bohm A. X-ray crystallographic and steady state fluorescence characterization of the protein dynamics of yeast polyadenylate polymerase. J Mol Biol. 2007; 366:1401–15. [PubMed: 17223131]
40. Yang ZR, Thomson R, McNeil P, Esnouf RM. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics. 2005; 21:3369–76. [PubMed: 15947016]
41. Batra VK, Beard WA, Shock DD, Krahn JM, Pedersen LC, Wilson SH. Magnesium-induced assembly of a complete DNA polymerase catalytic complex. Structure. 2006; 14:757–66. [PubMed: 16615916]
42. Steitz TA. DNA polymerases: structural diversity and common mechanisms. J Biol Chem. 1999; 274:17395–8. [PubMed: 10364165]

43. Lee SH, Choi HS, Kim H, Lee Y. ERK is a novel regulatory kinase for poly(A) polymerase. Nucleic Acids Res. 2008; 36:803–13. [PubMed: 18084034]
44. Shimazu T, Horinouchi S, Yoshida M. Multiple histone deacetylases and the CREB-binding protein regulate pre-mRNA 3′-end processing. J Biol Chem. 2007; 282:4470–8. [PubMed: 17172643]
45. Balbo PB, Meinke G, Bohm A. Kinetic studies of yeast polyA polymerase indicate an induced fit mechanism for nucleotide specificity. Biochemistry. 2005; 44:7777–86. [PubMed: 15909992]
46. Doublie S, Sawaya MR, Ellenberger T. An open and closed case for all polymerases. Structure. 1999; 7:R31–5. [PubMed: 10368292]
47. Doublie S, Ellenberger T. The mechanism of action of T7 DNA polymerase. Curr Opin Struct Biol. 1998; 8:704–12. [PubMed: 9914251]
48. Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Yamashita RA, Yin JJ, Zhang D, Bryant SH. CDD: a Conserved Domain Database for protein classification. Nucleic Acids Res. 2005; 33:D192–6. [PubMed: 15608175]
49. Geer LY, Domrachev M, Lipman DJ, Bryant SH. CDART: protein homology by domain architecture. Genome Res. 2002; 12:1619–23. [PubMed: 12368255]
50. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215:403–10. [PubMed: 2231712]
51. Galtier N, Gouy M, Gautier C. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput Appl Biosci. 1996; 12:543–8. [PubMed: 9021275]
52. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 2001; 17:754–5. [PubMed: 11524383]
53. Felsenstein, J. PHYLIP 3.6a2. University of Washington; 2001.
54. Martin G, Keller W. Mutational analysis of mammalian poly(A) polymerase identifies a region for primer binding and catalytic domain, homologous to the family X polymerases, and to other nucleotidyltransferases. EMBO J. 1996; 15:2593–603. [PubMed: 8665867]
55. Yang Q, Faucher F, Coseno M, Heckman J, Doublié S. Purification, crystallization and preliminary X-ray diffraction of a disulfide cross-linked complex between bovine poly(A) polymerase and a chemically modified 15-mer oligo(A) RNA. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2011; 67:241–4.
56. Otwinowski, Z.; Minor, W. Macromolecular Crystallography, part A. Vol. 276. Academic Press; 1997. Processing of X-ray Diffraction Data Collected in Oscillation Mode; p. 307-326.
57. Vagin A, Teplyakov A. MOLREP: an Automated Program for Molecular Replacement. Journal of Applied Crystallography. 1997; 30:1022–1025.
58. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010; 66:486–501. [PubMed: 20383002]
59. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr. 2002; 58:1948–54. [PubMed: 12393927]
60. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. 2010; 38(Suppl):W529–33. [PubMed: 20478830]
61. The PyMOL Molecular Graphics System, Version 1.5.0.4. Schrödinger, LLC;

Highlights

1. Three clades of a biochemically conserved poly(A) polymerase were identified in vertebrates.
2. The crystal structure of a first representative of the PAPγ clade was solved.
3. PAPγ closes around a bound 3′dATP and Ca^{2+}.
4. PAPs possess a conserved catalytic cavity, whereas surface residues are more divergent.

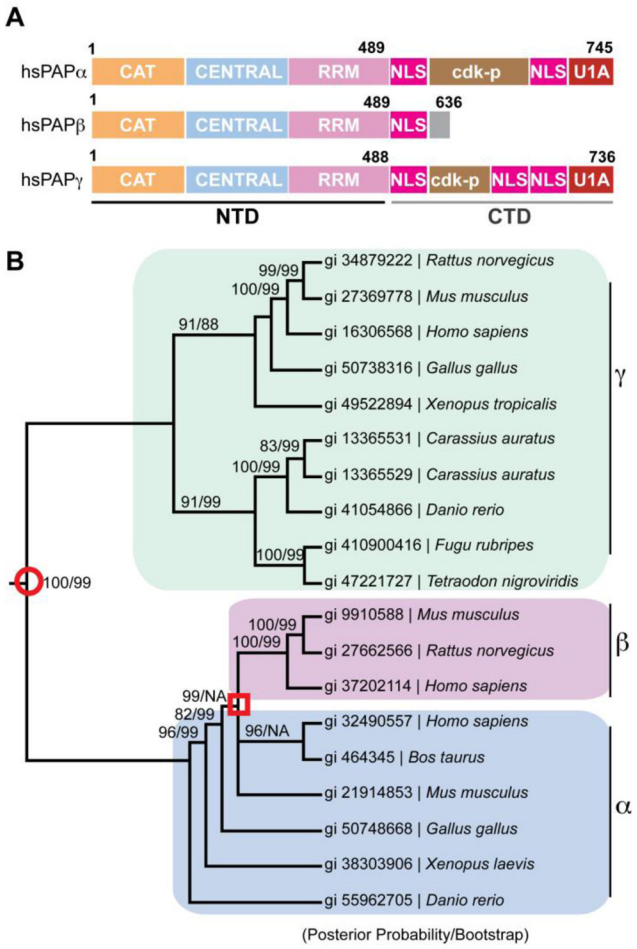

Figure 1. Poly(A) polymerases are grouped into three clades

(A) Domain organization of human poly(A) polymerases α, β, and γ. CAT is the abbreviation for the catalytic domain; RRM, RNA recognition motif; NLS, nuclear localization signal; cdk-p, cyclin-dependent kinase phosphorylation sites; U1A, U1A-interaction motif; NTD, N-terminal domain; CTD, C-terminal domain. The domain and motif sizes are not proportional to the length of the amino acid sequences. A detailed sequence alignment is shown in Figure S1.

(B) Phylogeny of vertebrate PAP sequences, constructed by MCMC simulation and the neighbor joining (NJ) algorithm (see METHODS for details). Statistical support for edges is described by MCMC posterior probabilities, followed by bootstrap support based on NJ calculations.

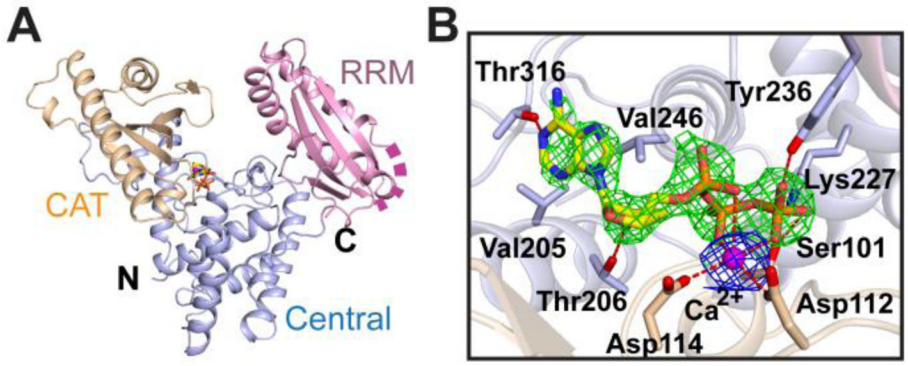

Figure 2. Structure of hPAPγ bound to 3′dATP and Ca^{2+}

(A) Human PAPγ is shown in cartoon representation and the color scheme is the same as in Figure 1A. The catalytic domain comprises residues 59-172, the central domain residues 17-58 & 173-351, and the RRM, residues 352-501. 3′dATP is shown as a stick model and Ca^{2+} as a sphere (magenta). (B) A close-up view of 3′dATP in the binding pocket of hPAPγ. Residues interacting with 3′dATP and Ca^{2+} are shown and colored according to the domain they belong to. Hydrogen bonds are represented by red dashed lines. An Fo−Fc omit map contoured at 3 σ (calculated before building 3′dATP in the map) is shown as a green mesh and a 4 σ anomalous difference Fourier map (blue mesh) overlays on top of the Ca^{2+} ion.

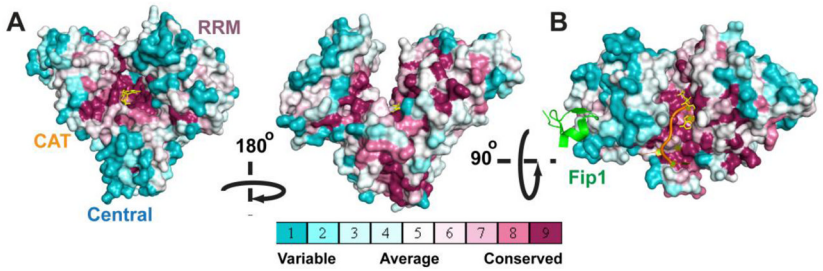

Figure 3. Conservation of the PAP N-terminal domain

(A) The conservation of PAP NTD (central, catalytic, and RRM domain) was calculated with the ConSurf server [60] and displayed with Pymol[61]. Conserved residues are shown in shades of magenta and variable residues in shades of blue. The bound 3′dATP is shown in yellow. (B) RNA (orange) from the Pap1-RNA complex (PDB: 2Q66 [28]) and Fip1 fragment (residues 80-105, green) from the yeast Pap1-Fip1 complex (PDB: 3C66 [21]) were overlaid on top of the PAP NTD (bottom panel) to illustrate the binding surfaces for RNA and Fip1.

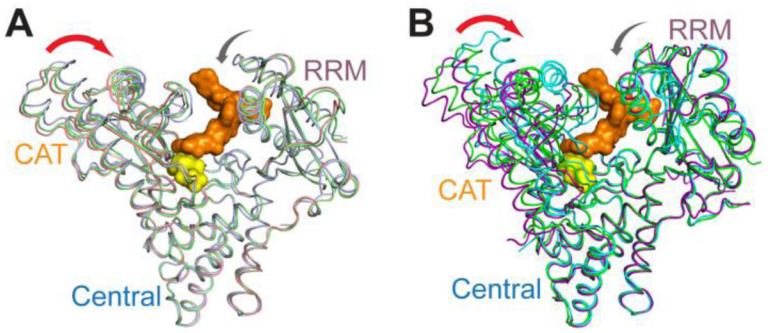

Figure 4. Superpositions of PAPs reveal domain movements

The structures were superimposed based on the central domain of hPAPγ (residues 17-58 & 173-351) to observe the movement of the catalytic domain and RRM.

(A) Superposition of three mammalian PAP structures (bPAPα-3′dATP-Mn^{2+} (PDB: 1Q79 [24]), pink; bPAPα-3′dATP-Mg^{2+} (PDB: 1Q78 [24]), green; and hPAPγ-3′dATP-Ca^{2+} (PDB:4LT6), lightblue) (B) Superposition of three yeast complexes (yPap1-3′dATP-Mn^{2+} (PDB: 1FA0 [26]), purple; yPap1- Mg^{2+} (PDB: 2HHP [39]), dark green; and yPap1-ATP-Mg^{2+}-RNA (PDB: 2Q66 [28]), cyan. ATP (yellow) and RNA (brown) from the yeast Pap1-RNA complex (PDB:2Q66 [28]) are modeled in the substrate binding cleft and shown in surface representation.