Published in final edited form as:

J Mol Biol. 2007 March 9; 366(5): 1401–1415.

X-ray Crystallographic and Steady State Fluorescence Characterization of the Protein Dynamics of Yeast Polyadenylate Polymerase

Paul B. Balbo, Joe Toth, and Andrew Bohm

*Tufts University School of Medicine and the Sackler School of Graduate Biomedical Sciences,
Department of Biochemistry, 136 Harrison Avenue, Boston, MA 02111*

Keywords

Poly(A) polymerase; induced fit mechanism; domain movement; crystal structure

Introduction

Poly(A) polymerase (PAP) catalyzes the 5'-to-3' polyadenylation of messenger RNA in a template-independent reaction^{1–4}. The synthesis of poly(A) tails occurs for all eukaryotic transcripts with the exception of the major metazoan histone mRNAs, and the poly(A) tail provides a universal handle by which transport and translation machinery can recognize and physically manipulate mRNAs⁵. PAP is a catalytic subunit of the cleavage/polyadenylation factor (CPF), a multiprotein complex that functions in the 3'-end processing of pre-mRNA in vivo^{6–10}. Although components of the CPF complex are required for various functions including the proper recognition of *cis*-elements of pre-mRNA, endonucleolytic cleavage at the poly(A) site, and regulation of poly(A) tail synthesis and termination^{6,11–19}, PAP retains adenylyltransferase activity and nucleotide specificity when separated from the complex²⁰. The activities of the CPF complex are conserved among eukaryotes⁵, and yeast PAP serves as an excellent model system for the study of template-independent polymerases.

Our previous structural studies²¹ have shown that yeast PAP is composed of three globular domains which are organized to form a central cleft (approximately 35 Å long and 25 Å deep) with the active site located at its base. The N-terminal (palm) domain is homologous to the nucleotide binding domains^{22,23} that are present in other members of the DNA polymerase β superfamily of enzymes, a group that includes DNA polymerase β , CCA adding enzymes, terminal deoxynucleotidyl transferases, 2'-5' synthetases, and many antibiotic nucleotidyltransferases²⁴. The N-terminal domain contains three conserved aspartic acid residues that directly coordinate two catalytically-required Mg^{2+} ions. One ion serves as a co-substrate ($MgATP^{2-}$) and the other is involved in coordination of both ATP (via a non-bridging α -phosphate oxygen) and the 3'-hydroxyl group of the poly(A) substrate. The reaction proceeds through an anionic phosphorane intermediate, and a major source of catalytic power in these enzymes is derived from the dispersion of this negative charge via the two metals^{25–29}. The middle domain is analogous to the “fingers” domain of other nucleic acid polymerases and contains catalytically important residues that interact with the phosphate groups of the

Corresponding author: Andrew Bohm; Tel: 617-636-2994, Fax: 617-636-2409, E-mail: andrew.bohm@tufts.edu

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

nucleotide (substrate) and pyrophosphate (product)^{22,23}. The C-terminal domain of PAP has no homolog in other polymerases and lies on the opposite side of the large central cleft from the palm domain. From chemical and UV-crosslinking experiments, the C-terminal domain of both yeast and bovine enzymes has been implicated in RNA binding^{30–32} and appears to contain an RNA recognition motif (RRM)^{21,33}. Additionally, the C-terminal domain has been shown to mediate protein-protein interactions with other polyadenylation factors by a direct interaction with Fip1 in yeast^{31,34} and, analogously, by formation of a ternary complex involving (human) hFip1 and CPSF-160 in mammals³⁵.

PAP has been thought to exhibit large scale rigid body domain movements based on the comparison of previous crystal structures of the enzyme²¹ and analogy to other nucleotidyltransferases^{36–38}. In the previous yeast PAP crystal structure²¹, the asymmetric unit cell contained two molecules that differed by 4.8° in the relative orientation of the N-terminal (palm) and middle domains. These crystals were prepared by soaking nucleotide and metal (3'dATP and Mn²⁺) into preformed crystals of the apo-enzyme. Two nucleotides were bound, one at the presumed location of the incoming ATP (the ATP subsite), and the other at the location of the 3' end of the mRNA prior to elongation (the poly(A) subsite). In these molecules, the active site, located at the interface between the N-terminal and middle domains, is in an open conformation and no enzyme-substrate contacts involving the adenine moiety of the incoming nucleotide are observed. Consequently, it is not possible to deduce the structural basis for nucleotide specificity from these structures. Additionally, only very small structural differences were observed between the domains of the empty and nucleotide-soaked molecules, thus there was no evidence for significant domain movement induced by the presence of the nucleotide. Structures of bovine PAP in the presence of 3'dATP with either Mn²⁺ or Mg²⁺ have also been determined^{33,39}. These structures were derived from the same crystal form and reveal a more open state of the enzyme, wherein the N- and C-terminal domains are rotated away from one another causing the central cleft to become wider than that seen in the yeast structures.

Other DNA polymerase β superfamily members exhibit conformational changes between open and closed states that involve rigid body domain movement, and these domain movements are believed to be integral to the induced fit mechanism by which the correct nucleotide is selected in these enzymes^{23,36–38}. For the template-dependent DNA and RNA polymerases, various methods including x-ray crystallography^{22,28}, steady-state⁴⁰ and stopped-flow fluorescence^{41,42}, and NMR⁴³ indicate that these polymerases are “open” when bound to primer and template DNA but “closed” around the nascent base-pair when the correct incoming nucleotide is bound. In these enzymes, the catalytic step occurs in the closed state while nucleotide binding and translocation of DNA/RNA occurs in the open form^{28,44}. Computational studies also support the idea that domain closure in these enzymes promotes catalysis by stabilizing the substrates, active site residues and required metal cofactors in an appropriate geometry²⁵.

As stated above, the open enzyme conformations observed in the available crystal structures of PAP do not exhibit enzyme-substrate interactions that account for nucleotide specificity^{21,33,39}. Therefore, it is reasonable to propose that proper nucleotide recognition and catalysis by PAP requires the enzyme to adopt a more closed conformation than those seen in the crystal structures reported to date. Although molecular dynamics simulations and normal mode analysis can suggest the possible motions available to a protein, it is difficult to determine which conformations are truly accessible to the molecule by computational means alone. Also, it is impossible to determine whether such movements are important for catalysis without experimentation. Furthermore, unlike the DNA or RNA polymerases that recognize correctly base-paired, duplex nucleic acid, PAP catalyzes adenylyltransfer in a template-independent manner. Consequently, the physiochemical mechanism for substrate specificity need not be

strictly conserved. To understand the molecular basis of PAP catalysis requires better structural and biochemical characterization of the enzyme-substrate complex in the closed conformation.

To address these issues, we report here the structure of two new crystal forms of yeast poly(A) polymerase as well as the results of steady state fluorescence studies designed to probe the extent and nature of PAP domain flexibility. One of the new structures reveals the enzyme in the most closed conformation seen to date. The second shows the enzyme in an intermediately closed state. Analysis of these along with our previous structures indicates that yeast poly(A) polymerase is capable of movement via two hinge regions and along two well-defined rotation axes. Modeling of the closed, substrate-bound state provides evidence for a number of specific interactions that form upon domain closure involving: (1) the triphosphate moiety of ATP (or pyrophosphate) and Lys215 and Tyr224, and (2) the adenine moiety of ATP and Asn226, suggesting a prominent role for this residue in determining nucleotide specificity. Finally, we present results from fluorescence studies that suggest that the closed enzyme conformation (with respect to the N-terminal and middle domains) is stabilized in the enzyme-substrate ternary complex (that corresponds to the Michaelis complex) and that this stabilization requires the molecular recognition of both the 6-amino group and triphosphate moiety of ATP.

Results

New crystal structures of poly(A) polymerase

To characterize the conformational flexibility of PAP, we have determined two new crystal structures of this enzyme. To facilitate our discussion, we will refer to the original structure²¹ as “crystal form 1,” and the new structures as “crystal form 2” and “crystal form 3.” Crystallographic parameters and statistics are presented in Table 1. Whereas crystal forms 1 and 2 each contain two copies of the polymerase in the asymmetric unit cell, crystal form 3 has only one. Thus these structures provide five independent views of the polymerase. For clarity, the individual molecules will be referred to as 1A, 1B, 2A, 2B, and 3. There is no obvious similarity in the molecular packing that gives rise to the three 3-D lattices, and the crystal contacts appear randomly distributed among the surface-accessible residues (supplemental material). As expected, residues within the interior of the three domains are consistently less flexible than those at the surface, but as evidenced by the distribution of B-factors, there are significant differences in the atomic-scale flexibility of surface-exposed, flexible loops among the five copies of PAP (supplemental material). Thus, there is no one “best” model of the molecule. In addition to the atomic-scale differences, there are significant domain-scale differences which, as discussed below, have lead us to conclude that the domains of PAP move along two well-defined trajectories.

Crystal form 2 is of space group $P2_12_12_1$, and contains two molecules in the asymmetric unit cell. These crystals are much more densely packed ($V_M = 2.4 \text{ \AA}^3/\text{Da}$) than crystal form 1 ($V_M = 4.1 \text{ \AA}^3/\text{Da}$) but only somewhat more densely packed than crystal form 3 ($V_M = 2.5 \text{ \AA}^3/\text{Da}$). In spite of their close packing the two molecules within the asymmetric unit of crystal form 2 represent intermediately closed states relative to the molecules seen in the other two crystal forms.

Crystal form 3 belongs to space group $P2_1$ and, at 1.8 Å resolution, provides a significantly clearer picture of the yeast enzyme than that which was reported earlier. In particular, many side chains with ambiguous electron density in the earlier structure are much improved. These crystals have one molecule in the asymmetric unit cell, and this structure represents the most closed state of the enzyme yet observed. Crystals in this space group are originally grown in the presence of AMP-CPP ($K_d = 75 \text{ }\mu\text{M}$)⁴⁵, but the electron density within the active site was difficult to interpret and crystals grew equally well in the absence of nucleotide analogs as in their presence. At high citrate concentrations, a well ordered citrate molecule along with a

single Mg^{2+} ion (carried over from the storage buffer) occupy the active site, but at lower concentrations the electron density within the active site, though still present, is difficult to model. Attempts to replace the citrate at the active site with various nucleotides and nucleotide analogs after crystals were soaked to remove the citrate were unsuccessful. In addition to the molecule at the active site, electron density for an additional citrate molecule was discovered. This lies at the interface between adjacent PAP molecules, perhaps explaining why citrate was absolutely required for growth of this crystal form. Additionally, kinetics experiments showed Mg-citrate to be a very weak inhibitor of polyadenylation ($\text{IC}_{50} \approx 20 \text{ mM}$, data not shown).

As noted above, PAP is composed of three globular domains which surround a large, central cleft. Since the average movements between any given pair of domains in our five structures is only $\sim 4^\circ$, we used the program ESCET^{46,47} to help us define which parts of the molecule act as rigid units. Error-scaled difference distance matrices for the ten possible combinations of structures were calculated, and as shown in Figure 1, the composite of these ten matrices clearly shows that PAP is composed of three rigid units. The precise domain boundaries were defined based on examination of the difference distance matrices, and by repeated structural superpositions followed by manual inspection of the resulting overlaps. These analyses lead us to define the domain boundaries as follows: Residues 40-190 belong to the catalytic palm domain. Residues 4-39 and 191-353 belong to the Middle domain. Residues beyond 353 belong to the C-terminal domain.

To visualize the domain movements, the middle domains of the five molecules were superimposed, and the rotations and translations required to align either the N- or C-terminal domains were calculated (Figure 2, Table 2). The structures of the individual domains are very similar in the five structures with RMS deviations in alpha carbon positions ranging from 0.3 to 1.8 Å. The analysis suggests that PAP can move about two well-defined rotation axes: one between the N-terminal and middle domains, the other between the middle and C-terminal domains. Though the degree of movement varies, all of the structures are generally consistent with movement about the same two rotation axes. The next sections describe each of these domain movements in more detail.

Movement of the N-terminal (palm) domain

The N-terminal domain (residues 40-190) moves perpendicular to the cleft, and movements of this domain directly result in widening or narrowing of the substrate binding pocket at the interface of the N-terminal and middle domains (Figure 2C). The largest palm movements are those between structure 3 and structure 1A. Comparison of these structures shows a 8.6° rotation of this domain with a small translational component (0.13 Å). Translational components among other yeast structures along this axis were also quite small, and thus these movements are well described by a simple rotation. The hinge regions between N-terminal and middle domains are composed of two structural elements. The first is a loop (residues 37-41) between helix A, which moves as part of the middle domain, and helix B, which is part of the N-terminal domain. The second is near the middle of helix F, which has a slight kink at Gly190. Kinked helices function as hinges in other proteins⁴⁸, but the location of this hinge, at an amino acid mostly buried within the interior of the protein, and in the middle of a helix spanning two domains, is notable. A number of side chains are repositioned at the interface of the N-terminal and middle domains upon domain movement; Met310 and His314, located adjacent to one another at the base of the cleft on the back side, undergo particularly large side chain movements. There are also significant changes in both the side chain and backbone positions in the loop between 278 and 283. This loop is located well outside the central cleft on the outside surface of the N-terminal domain. The loop appears to slide along the surface of this domain as part of the conformational rearrangement. Sliding motions of this sort have been observed in a number of other proteins and have been reviewed⁴⁸. This movement is potentially

interesting since the nearby N-terminal region of PAP is thought to interact with an as yet uncharacterized subunit of the cleavage/polyadenylation complex, and thus could function in regulating mRNA 3'-end formation³².

The most functionally significant consequence of the movement between the N-terminal and middle domains is a narrowing of the active site cleft. Though the structure of the most closed state, crystal form 3, does not contain a nucleotide at the active site, there is ample space for ATP and the 3' terminus of the poly(A) substrate. To examine the consequences of domain closure on substrate binding we modeled the closed, substrate-bound state by superimposing the N-terminal domain of molecule 1A onto the equivalent domain of molecule 3, and then displayed the structure of crystal form 3 along with the nucleotides (3'-dATP and 3'-dAMP) from 1A. The predicted structure is shown in Figure 3A and depicted schematically in 3B. Because of the previously noted agreement between the metal-nucleotide conformations observed in crystal structures of yeast PAP²¹ and other nucleotidyltransferases²⁸ and the fact that the N-terminal domain is essentially identical in all of our crystal structures, we assert the nucleotide positioned in this way should be a reasonable estimate of the nucleotide-bound state of the most closed structure, molecule 3. The nucleotide is directly bound to the N-terminal domain by interactions between the Mg-triphosphate and the conserved acidic residues, Asp100, Asp102, and Asp154. Therefore, our analysis identifies residues of the middle domain that are moved closer to the nucleotide as a result of domain closure.

The nucleotide modeled into molecule 3 is significantly closer to residues of the middle domain than it is in the most open structure (molecule 1A). The residues Lys215, Tyr224, and Asn226 are part of Helix G and the loop immediately C-terminal to it. To a close approximation, these structural elements move as a rigid body as a consequence of the deformation of Helix F that contains Gly190. This is summarized in the schematic representation in Figure 3B. Specifically, the C α of Lys215 and Tyr224 (respectively located on and immediately adjacent to Helix G of the middle domain) are translated by approximately 2.5 Å relative to their position in 1A, the most open structure. This results in movement of the side chains of these residues such that the Lys-N ϵ and Tyr-OH are both translated by ~1 Å towards the non-bridging oxygens of the γ - and β -phosphates of MgATP²⁻, respectively (Figure 3A). The proximity of these residues to the triphosphate group of ATP was previously noted^{21,33,39}, but the precise interactions could not be deduced because of the distances between these residues and the substrate in the open conformation seen in these structures. It should be noted that the α -phosphate oxygens appear to lack coordination with any active site residues. Additionally, Asn226 is located 3.0 Å from the adenine base of MgATP²⁻ in molecule 3, suggesting that it is available to H-bond to the substrate via interaction with possibly the 6-amino or N7 groups of the substrate. Mutagenesis studies⁴⁹ have also implicated N226 in RNA binding and, as discussed later, we cannot rule out the possibility that this side chain serves a dual function. We emphasize that the most closed structural model does not exhibit contact between active site residues and the adenine base moiety of the nucleotide that would constitute molecular recognition of the substrate. Nonetheless, based on this model, it is reasonable to deduce that further motion along the trajectory described above would allow these residues to directly contact the ATP substrate during catalysis.

Movement of the C-terminal domain

In contrast to the N-terminal domain movements, movements of the C-terminal domain are roughly parallel, not perpendicular, to the axis of the cleft (Figure 2B,C). Consequently, these movements do not significantly change the overall width of the cleft. Molecules 3 and 2A differ most in this region. The C-terminal domains of these structures exhibit a 7.1° rotation as well as a 0.63 Å translation. Here, only one region of the polypeptide chain serves as a hinge and is located at the N-terminus of helix M, near Asp353. The region just after the hinge is held

tightly to the remainder of the C-terminal domain through interactions between hydrophobic residues F354, F355, F405, F409, Y465, and V521. There are fewer noncovalent interactions between the C-terminal and middle domains than between the N-terminal and middle domains. There are also fewer rearrangements resulting from the domain rotation. Based on the positions of the incoming nucleotides observed crystallographically^{21,39} and modeled in Figure 3, the C-terminal domain appears to be too far from the active site to participate in chemistry. Nonetheless, movement of the C-terminal domain alters the position of residues across the central cleft from the active site and could affect the access and egress of substrates. (The path of the poly(A) substrate is unknown, but could potentially lead up or backwards out of the central cleft.) Additionally, displacement of the C-terminal domain would be required to avoid clashing with the N-terminal domain upon formation of the closed conformation described above. Finally, Helix N, which lies across the central cleft, contains some of the only residues within the C-terminal domain which are conserved across the phyla. Thus the observed movements of the C-terminal domain are of potential import.

Comparison of the domain orientations observed in the yeast PAP structures to those of bovine PAP

As indicated earlier, bovine PAP differs significantly in sequence from the yeast enzyme. The two proteins share ~48% identity in the first two domains, but are substantially more divergent (less than 20% identity) in the 3rd domain. In addition, the mammalian homolog contains a 4th domain which is not present in the yeast enzyme. Despite these differences in sequence, the first three domains of bovine PAP adopt the same overall fold as the yeast enzyme. Bovine PAP has been solved in complex with a variety of different substrates in the active site, but all of the bovine structures are from the same crystal form. There are ~2.5° differences in the domain orientations in some of the bovine structures^{33,39}, but since these differences are small relative to the differences seen between the structures of the yeast enzyme we have presented a comparison of the yeast PAP structures with just one of the bovine structures (pdb code 1F5A) in Table 2. As mentioned above, the bovine PAP structure assumes a more-open conformation than does the yeast enzyme. Superimposition of the middle domains of PAP from the two species and examination of the relative positions of the N-terminal (palm) and C-terminal domains reinforces the conclusions presented above; namely that PAP is capable of movement along two well defined axes. The conservation of these axes of motion across species, despite significant sequence divergence, support the supposition that the movements we observe make an important contribution to the enzymatic function of all poly(A) polymerases.

Steady state fluorescence studies using A₃(2AP)A and A₄(2AP) as active site probes

As described above, the movement of the N-terminal and middle domains alters the active site geometry. Steady state fluorescence experiments were performed that utilized fluorescent substrate analogs as molecular probes in order to assess the solvent environment of the active site of PAP under various conditions where enzyme-substrate complexes form. The degree of solvent exposure or protection experienced by the active site-bound probes was taken as evidence of the enzyme assuming an open or closed domain conformation, respectively, under that condition. The probes employed were pentameric polyadenylate (A₅) molecules labeled with 2-aminopurine (2AP) at either the 3'-terminal or penultimate position: A₄(2AP) or A₃(2AP)A, respectively. The fluorescence properties of 2-aminopurine-labeled nucleotides and nucleic acids have been successfully used previously to characterize protein-nucleotide and nucleoprotein systems^{41,42,50}, and in our study, these reagents enabled the survey of three specific locations in the active site.

The fluorescence emission spectra of A₃(2AP)A and A₄(2AP) were measured either alone or in the presence of excess PAP with or without MgATP²⁻, MgCTP²⁻, or MgAMP-CPP²⁻, and the results are shown in Figure 4. For both the probes A₃(2AP)A and A₄(2AP), the fluorescence

emission intensity increased upon binding the enzyme with no change in the emission wavelength maximum. This result is consistent with general solvent effects^{*51–53} and suggests that the probe experiences a different solution environment when present in enzyme-substrate complexes. The maximum fluorescence change was observed in the presence of MgATP^{2-} . Formation of either a $\text{PAP} \cdot \text{A}_3(2\text{AP})\text{A}$ or $\text{PAP} \cdot \text{A}_4(2\text{AP})$ binary complex resulted in an intermediate fluorescence increase which was not affected by the presence of MgAMP-CPP^{2-} or MgCTP^{2-} , except in the case of MgCTP^{2-} and $\text{A}_4(2\text{AP})^\dagger$. This result suggests that the ternary complexes involving PAP, either A_5 analog, and specifically MgATP^{2-} exhibits further changes in the solvent environment of the active site.

In order to better characterize the interaction between the polyadenylate probes and the enzyme, the probe was titrated with PAP in the absence or presence of an excess amount of MgATP^{2-} . The results of the titration of $\text{A}_3(2\text{AP})\text{A}$ are shown in Figure 5. This experiment measures the dissociation constant of the polyadenylate probe from the enzyme-substrate binary or ternary complex and the maximum change in fluorescence observed upon complex formation. These are described by Equations 1 and 2 (below), respectively, along with the K_d values determined from this experiment.

$$\text{PAP} \cdot \text{A}_3(2\text{AP})\text{A} = \text{PAP} + \text{A}_3(2\text{AP})\text{A} \quad K_1 = 52.1 \pm 8.8 \mu\text{M} \quad [1]$$

$$\text{PAP} \cdot \text{MgATP}^{2-} \cdot \text{A}_3(2\text{AP})\text{A} = \text{PAP} \cdot \text{MgATP}^{2-} + \text{A}_3(2\text{AP})\text{A} \quad K_2 = 16.4 \pm 3.3 \mu\text{M} \quad [2]$$

The observation that the probe titrates to a different end-point at saturating enzyme concentrations with or without MgATP^{2-} suggests that the probe experiences (on average) different environments in the binary and ternary complex. The dissociation constants K_1 and K_2 can be related to the kinetic parameters K_{ia} and K_a , respectively (for description of the kinetic parameters and mechanism of PAP, see Balbo et. al. (2005)⁴⁵). As was previously measured⁴⁵ for the substrate A_{18} , $K_{ia} = 93.2 \pm 28.9 \mu\text{M}$ and $K_a = 46.8 \pm 12.4 \mu\text{M}$. Both of these results indicate that the polyadenylate substrate binds to the enzyme with modestly higher affinity when MgATP^{2-} is also present, and this agreement supports the validity of the use of these probes for studying PAP-substrate interactions.

To investigate the solvent accessibility of the active site, acrylamide quenching experiments were performed using $\text{A}_3(2\text{AP})\text{A}$ or $\text{A}_4(2\text{AP})$ in the absence or presence of PAP and various nucleotides, and the results are shown in Figure 6. If formation of an enzyme-substrate complex stabilized the closed conformation, protection of the fluorescent probe from acrylamide would be expected, provided that solvent was excluded from the active site in the closed state. The use of both probes gives insight into the solvent environment at both the 3'-terminal and penultimate residue positions of polyadenylate. Formation of either $\text{PAP} \cdot \text{RNA}$ binary complex resulted in decreased acrylamide quenching efficiency. The observed solvent protection and increased fluorescence emission observed for the binary complexes may simply be a consequence of adsorption of nucleic acid probe to the active site surface rather than an indication of domain closure. Samples containing MgAMP-CPP^{2-} exhibited quenching behavior similar to the PAP binary complexes suggesting that this analog did not induce any

*The fluorescence of 2AP is sensitive to solvent polarity. Additionally, when 2AP is incorporated in single stranded polyadenylate, fluorescence quenching occurs via base stacking. Thus, we considered that the increase in fluorescence was the result of a decrease in base-stacking interactions. Although distortion of base-stacking interactions upon enzyme substrate complex formation cannot be entirely ruled out, we deduce that solvent effects are primarily responsible for the observed fluorescence increase. This is based on similar fluorescence changes observed in control experiments for both the labeled RNAs and the 2AP ribonucleotide measured in water or water/alcohol mixtures (not shown).

†In the case of $\text{A}_4(2\text{AP})$ and CTP, lower fluorescence emission intensity was observed (Figure 4B). We speculate this could be due to quenching, mediated by the cytidine base, of the 2AP label by that would be exposed at the 3'-terminal position. Previous studies indicate that a 2AP base adjacent to a cytidyl residue in nucleic acid is quenched due to collision of the bases (see reference⁵⁰).

further change in the active site. The quenching results for $\text{PAP} \cdot \text{MgCTP}^{2-} \cdot \text{A}_4(2\text{AP})$ were similar to the $\text{PAP} \cdot \text{A}_4(2\text{AP})$ binary complex ($K_{SV} = 3.16$ and 3.30 M^{-1} , respectively), but the results for $\text{PAP} \cdot \text{MgCTP}^{2-} \cdot \text{A}_3(2\text{AP})\text{A}$ resembled that of free $\text{A}_3(2\text{AP})\text{A}$ ($K_{SV} = 2.96$ and 3.06 M^{-1} , respectively). This could indicate that for the enzyme-substrate ternary complex involving the incorrect nucleotide, CTP, the poly(A) substrate is solvated to a different extent at either of the terminal or penultimate position surveyed by the probes. Alternatively, the anomalous result with CTP could arise from a combination of factors or reflect substrate binding in an alternate mode. Most significantly, when MgATP^{2-} was present, acrylamide quenching was substantially less efficient regardless of which probe was used. This result suggests that the active site in the enzyme-substrate ternary complex with MgATP^{2-} is more solvent-protected, and possibly reflects domain closure.

An analogous set of experiments was performed employing 2-aminopurine riboside triphosphate (2AP) as an ATP analog. Kinetic experiments indicated the K_m (here, $K_m = K_b$, defined in reference ⁴⁵) for this alternative substrate is $19.7 \pm 2.5 \mu\text{M}$ (data not shown), comparable to the previously measured value⁴⁵ for ATP ($35.9 \pm 12.6 \mu\text{M}$). Binding of $\text{Mg}(2\text{AP})^{2-}$ ($1.3 \mu\text{M}$) to PAP ($61 \mu\text{M}$) in the presence or absence of A_5 ($64 \mu\text{M}$) resulted in a very small change in 2AP fluorescence emission ($\pm 6\%$ relative to 2AP alone). Furthermore, although formation of either a $\text{PAP} \cdot \text{Mg}(2\text{AP})$ binary or $\text{PAP} \cdot \text{A}_5 \cdot \text{Mg}(2\text{AP})$ ternary complex did result in protection from acrylamide quenching, the quenching efficiency of either of these complexes was very similar. The reduced quenching efficiency observed in these complexes relative to 2AP alone is most likely due to adsorption to the active site surface and does not reflect domain closure. Additionally, the degree of solvent protection observed for both of these complexes ($K_{SV} = 4.8 \text{ M}^{-1}$, data not shown) was not as great as those observed for the ternary complexes involving the 2AP-labeled RNAs and MgATP^{2-} ($K_{SV} = 1.4\text{--}1.7 \text{ M}^{-1}$, Figure 6). Taken together, these results suggest that the $\text{Mg}(2\text{AP})^{2-}$ ribonucleotide base is relatively exposed to solvent in both the binary and ternary complexes and does not induce domain closure as was observed for MgATP^{2-} . This result further suggests the importance of the exocyclic 6-amino group of the ATP substrate in the induction of this effect, possibly by the specific recognition of the 6-amino group in the closed conformation state of the enzyme.

In summary, the fluorescence data together indicate that binding and recognition of both substrates together by PAP results in specific changes in the active site environment including decreased solvent accessibility. This suggests that the enzyme-substrate ternary complex is stabilized in a more-closed conformational state. Furthermore, the results of experiments utilizing $\text{Mg}(2\text{AP})^{2-}$, MgCTP^{2-} and $\text{MgAMP} \cdot \text{CPP}^{2-}$ suggest that this effect is specific for MgATP^{2-} and requires molecular recognition of the triphosphate and 6-amino moieties of the nucleotide.

Discussion

Rigid body domain movements of poly(A) polymerase

Poly(A) polymerase is composed of three globular domains^{21,33}. Here we describe the rigid body motion of the N-terminal (palm) and C-terminal domains of PAP, relative to the middle domain, as deduced from several crystal structures. The active site is composed of residues from both the N-terminal and middle domains, and the substrates bind at the interface, thus the motion associated with these domains certainly impacts the catalytic activity of the enzyme. This notion is supported by the previous observation that nucleotides bound at the active site in the open conformation show no specific contacts between active site residues and the adenine base that could account for substrate specificity^{21,33,39}. The analysis presented in this work allows us to predict additional, specific enzyme-substrate contacts that form in the closed state. The purpose of the fluorescence experiments described herein was to investigate the solvent accessibility of the active site in the presence of substrates and substrate analogs. These

experiments reported on the open or closed conformational state of the N-terminal and middle domains of the enzyme that predominates under each condition. The results suggest that the closed domain conformation is stabilized upon formation of the $E \cdot A_n \cdot \text{MgATP}^{2-}$ ternary complex, which corresponds to the Michaelis complex.

The C-terminal domain has been demonstrated to be important in protein-protein interactions³². Specifically, PAP binds tightly to Fip1 which, in turn, interacts with other proteins in the CPF complex and so is thought to mediate the formation of larger protein subcomplexes within the CPF. Notably, Fip1 binds Yth1, a Zn finger-containing polyadenylation factor that binds pre-mRNA near the poly(A) site of endonucleolytic cleavage³⁴. Fip1 also binds Rna15 (active as an Rna14/Rna15 complex) which interacts with AU-rich regions upstream of the poly(A) site⁵⁴. This network of interactions is conserved in higher eukaryotes, including humans^{55,56}. A specific role for domain movement involving the middle and C-terminal domains has not been previously proposed. However, due to its association with Fip1, it is tempting to speculate a role in enzyme regulation. Experimental verification of such a role will require a more thorough biochemical description of the interaction of regulatory CPF subunits with PAP/Fip and their structural organization.

Effects of domain movements in PAP on chemistry and substrate specificity: an induced fit mechanism

The structural determinants of ATP binding and nucleotide specificity have not been well understood because of previous observations of the enzyme in an open conformation in crystal structures and uncertainty in the conformation of the adenine base of the bound nucleotide in the closed conformation^{21,33,57}. In these structures, the bound nucleotides (both at the ATP and poly(A) subsites) are in direct contact with residues in the N-terminal domain; the former interaction is primarily via the Mg-triphosphate moiety. In the analysis described in the present work, we superimpose structures 1A and 3 (Figure 3) in order to better characterize changes in enzyme-substrate interactions that occur when the enzyme assumes the closed conformation.

Firstly, the analysis supports the proposal by Bard et al²¹ that Asn226 is oriented such that it is capable to interact with ATP, thus suggesting, in part, a structural rationale for nucleotide specificity. This residue clearly comes into closer proximity of the adenine of ATP in our model as a consequence of domain movement and appears to be poised to participate in H-bonds with this moiety, possibly via the exocyclic 6-amino or endocyclic N7 group of the base. Martin et al^{33,39} noted that Asn226 (bovine Asn239) was 8Å away from the nucleotide in the bovine PAP structure and postulated that this residue was important in poly(A) binding, rather than nucleotide binding. Those authors also cite a mutagenesis study of yeast PAP⁴⁹ as supporting evidence, however that study did not involve a complete kinetic characterization of the mutant, thus the role of Asn226 in rate enhancement or on any particular kinetic parameter remains uncertain. The authors also proposed that Asn189 and Thr304 (bovine residues N202 and T317) play a prominent role in nucleotide selection based on crystallographic and mutagenesis data, and provided evidence that these residues interact with N1 of the adenine of ATP. Asn189 is part of the N-terminal domain, thus its position relative to MgATP^{2-} does not change upon domain closure, although Thr304, in the middle domain, moves 0.52Å toward the Asn189 in our comparison of molecules 1A and 3 (not shown). However, in either molecule 1A or 3 of yeast PAP, the side chain of Asn189 is 6.40Å away from the N1 group of ATP (not shown), which is in disagreement with the proposed interaction with N1. The discrepancy between these two proposals could be related to the difference in nucleotide conformation seen in the bovine and yeast crystal structures. The conformation of the incoming nucleotide in the yeast structure²¹, particularly the positions of the phosphorus atoms of the triphosphate moiety, the two metals, and catalytic aspartic acids, coincides remarkably well with these groups of the

DNA pol β structure²⁸, which supports the conformation of the nucleotide reported in molecule 1A²¹ and the model in Figure 3.

Secondly, the analysis supports that formation of the closed state brings the active site residues Tyr224 and Lys215, located in the middle domain, into proximity of the non-bridging oxygens of the β - and γ -phosphates of ATP, respectively. Because of their location at the active site, these residues have been previously implicated to interact with the triphosphate moiety of ATP in the substrate ternary complex and pyrophosphate in the product ternary complex^{21,33}, and mutation of the corresponding residues in bovine PAP resulted in decreased catalytic activity⁵⁸. Similar residues occur in the homologous enzyme, DNA polymerase β , as observed in the X-ray crystal structure of the ternary complex containing gapped DNA and (correctly base-paired) ddCTP²⁸. The side chains of two residues, Arg183 and Ser180, contact the β - and γ -non-bridging phosphate oxygens of the nucleotide, respectively, in the closed complex. The authors proposed that these function to stabilize the Mg-ddNTP in a chemically competent binding position. This was subsequently supported by biochemical studies⁵⁹. Lys215 and Tyr224 could serve the same function in PAP.

Finally, the analysis suggests that there are no protein residues participating in charge neutralization of either the 3'-OH or non-bridging oxygen of the α -phosphate of ATP. This is significant because during the reaction, the α -phosphate proceeds through a pentavalent phosphorane structure in the transition state upon nucleophilic attack by the activated (i.e., deprotonated) 3'-OH of poly(A), resulting in the generation of a negative charge at the transition state. In the chemical mechanism proposed for the DNA pol β -like enzymes, the α -phosphate non-bridging oxygen participates in coordination bonds with both metal ions, which are employed in charge neutralization. The dispersal of this charge is a major source of catalytic power in the polymerases that require two metals^{26–29}, and some one-metal nucleotidyltransferases have protein side chains that interact with the α -phosphate, presumably functioning in charge dispersal at the transition state⁶⁰. In the active site of PAP, there is cavity present in this region near the α -phosphate, consistent with other well studied two-metal polymerases.

To summarize, based on the observations that, in the closed structure modeled in Figure 3, the distances between the (i) the adenine of ATP and N226 and (ii) K215 and Y224 and the β - and γ -phosphate oxygens, respectively, are all still longer than ideal H-bond distances, we conclude that further motion of the N-terminal domain along the trajectory described here is required to achieve the fully closed, active state of the enzyme. Furthermore, steric considerations demand that this motion would require further displacement of the C-terminal domain.

Summary and conclusions

A comprehensive picture of the molecular mechanism of polyadenylate polymerase is beginning to emerge from this and recent kinetic, structural and biophysical studies of both yeast and bovine enzymes^{21,30,33,45,57,58}. Recent kinetic studies have revealed that PAP uses an induced fit mechanism for nucleotide specificity in which binding free energy from recognition of the correct substrate, MgATP²⁻, is used to lower the energy barrier of the rate-determining, central step (i.e. conversion of enzyme-bound substrates and products) by a ground state destabilization mechanism⁴⁵. The fluorescence data showing that MgATP²⁻, but not MgCTP²⁻, promotes the protection of the active site from solvent (interpreted here as stabilization of the closed enzyme conformation) suggests that the domain closure is an important feature of the induced fit mechanism of PAP. Furthermore, the substrate binding mechanism was also demonstrated to be in rapid equilibrium relative to the central step⁴⁵. This indicates that turnover is not limited by slow protein dynamics that would prevent substrate binding or dissociation.

Together, these data suggest a simple mechanism for the catalytic cycle of PAP. Firstly, substrates and products can rapidly associate with or dissociate from an open conformation of the enzyme. We propose that the open conformation allows for sampling of nucleotide substrates and facile dissociation of incorrect substrates which fail to induce the closed enzyme conformation. This is supported by the fluorescence quenching experiment with CTP reported here and consistent with previous work demonstrating that CTP is utilized as a substrate less efficiently than ATP^{31,45} and that cytidyltransfer is not promoted by ground state destabilization⁴⁵. In the Michaelis complex in the (forward) polyadenylation direction, a closed enzyme conformation is stabilized upon proper recognition of the correct substrates, thereby enabling efficient catalysis. Previous kinetic studies indicated that product dissociation from the enzyme was promoted by the relatively high K_m values, K_p and K_a' , for $MgPP_i^{2-}$ and poly(A), respectively⁴⁵. It remains to be determined if the enzyme product ternary complex is stabilized in an open or closed conformation, though the steady state kinetic data suggest the former. Rapid domain opening upon product formation would promote poly(A) dissociation in the product binding mode and allow it to rebind in the substrate mode for the next round of adenylyltransfer.

Methods and Materials

Chemicals

Reagents were of analytical grade. ATP, CTP and α,β -methylene-ATP (AMP-CPP) were from Sigma-Aldrich. 2-aminopurine riboside triphosphate (2AP) was from IBA Biologics (Göttingen, Germany). The oligonucleotide A₅ was purchased from TriLink (San Diego, CA); the fluorescent pentanucleotides A₃(2AP)A and A₄(2AP) were from Dharmacon (Lafayette, CO).

Protein purification

The enzyme used for all the crystallography and biochemical experiments was a recombinant yeast PAP that was constructed with a 32 amino acid C-terminal truncation and a C-terminal His₆ tag⁴⁵ and is referred to simply as PAP throughout this manuscript. These modifications do not affect the enzymatic properties of the polymerase³². PAP was purified by both Ni-affinity and ion exchange chromatography as described⁴⁵. Protein used for crystallization was concentrated to 20–30 mg/ml, exchanged into 20 mM phosphate pH 6.8, 20% glycerol, 2 mM MgCl₂, frozen in liquid nitrogen, and stored at –80°C prior to crystallization. Protein used for fluorescence studies was prepared and stored similarly, except MgCl₂ was omitted.

Crystallization and Structure Determination

Orthorhombic crystals (crystal form 2) of space group P2₁2₁2₁ were grown using the hanging drop method at 4°C by mixing 1 μ L of the thawed protein with 1 μ L of reservoir solution (20% PEG 8000, 100 mM magnesium acetate, 100 mM imidazole, pH 6.2 and 3% ethylene glycol). Crystals appeared after two weeks and continued to grow for approximately 1 additional week. Prior to data collection, crystals were transferred into a solution containing 20% ethylene glycol and 80% reservoir solution and flash frozen in liquid nitrogen.

Monoclinic crystals (crystal form 3) of space group P2₁ were originally grown at 4°C by combining 2 μ L of the thawed protein with 2 μ L of 20% isopropanol, 20% PEG 4000, 100 mM citrate pH 5.6 and 0.1% β -mercaptoethanol. However, hanging drops with reservoir seldom produced usable crystals. Crystals were consistently produced from streak-seeded hanging drop setups without reservoir (and omitting the 20% isopropanol). Crystals were transferred into a cryoprotectant solution containing either mother liquor and 25% ethylene glycol or 22% PEG 400 prior to data collection, and frozen in liquid nitrogen or gas stream at 100K.

Preliminary data were collected on an Xcaliber PX-ultra X-ray diffraction system at 100K (Oxford Diffraction). Higher resolution data from the $P2_1$ and $P2_12_12_1$ crystal forms were then collected at NSLS beamline X25 and X29, respectively. Diffraction data were processed with MOSFLM⁶¹ and HKL2000⁶² and molecular replacements were performed with CNS⁶³ and CCP4⁶⁴ using the structure of yeast PAP as the search model. The structures were traced with COOT⁶⁵ and subjected to simulated annealing refinement in CNS. Final stages of refinement were performed with REFMAC⁶⁶. For both structures, each PAP molecule was refined with three TLS groups with each domain corresponding to a group.

Steady state fluorescence measurements

Steady state fluorescence experiments were performed that utilized fluorescent substrate analogs as molecular probes. Specifically, the probes were pentameric polyadenylate (A_5) analogs that incorporated a fluorescent 2-aminopurine (2AP) base (an analog of adenine) at either the 3'-terminal, $A_4(2AP)$, or penultimate, $A_3(2AP)A$, position. Additionally, the nucleotide 2-aminopurine riboside triphosphate was used as an ATP analog.

Steady state fluorescence emission spectra were measured at room temperature using a Perkin Elmer LS-50B spectrofluorometer with a xenon flash tube light source. The instrument automatically performed dark current correction and corrected the block averaged signals from the sample and reference channels for drift in the spectral response of the photomultiplier tubes and transmission response of the beamsplitter by comparison to an internal rhodamine standard. Sample volumes of 180 μ L were measured in a 3×3 mm cuvette. The 2AP-labeled probes were excited at 315 nm and had an emission maximum at ~ 370 nm; absorbance from the sample at these wavelengths was negligible. Emission spectra were collected (330–450 nm) at a scan speed of 450 nm/min and three consecutive scans were averaged. The excitation and emission slit widths were 6.2 nm. In all cases, samples were corrected for background by subtracting the fluorescence of an identically prepared sample that contained all sample components minus the 2AP-labeled probe.

The buffer used in the measurement of fluorescence spectra, including the acrylamide quenching experiments, was 40 mM bis-Tris, pH 7.0, 20 mM NaCl, 10% glycerol, 0.18% reduced Triton X-100, and 10 mM β -mercaptoethanol. The concentrations of the reagents were: 105 μ M PAP (6.6 mg/mL), 1.91 μ M $A_4(2AP)$, 1.96 μ M $A_3(2AP)A$, and 1.36 μ M 2AP riboside triphosphate, where applicable. In experiments that included non-fluorescent nucleotides/analogs, their concentrations were 520 μ M in MgNTP (where NTP = ATP, CTP, or AMP-CPP). In order to minimize catalytic turnover during the fluorescence experiments, the concentration of free Mg^{2+} in the samples was made negligible by having a small molar excess (1.3-fold) of free nucleotide over $MgCl_2$ (free $[Mg^{2+}]$ calculated to be 10–20 μ M; the variation is due to small differences in actual nucleotide concentrations). Because PAP requires a second, catalytic Mg^{2+} ($K_d \approx 1.5$ mM)⁴⁵, Mg^{2+} -depletion renders the enzyme inactive. This method of metal depletion has previously been employed in the study of DNA polymerase β ; that enzyme had been shown capable of forming enzyme-substrate complexes in the absence of free metal that were competent for reaction upon the addition of metal⁶⁷.

The binding of the 2AP-labeled polyadenylate probes with PAP was investigated in titration experiments where the probe was held constant (1.95 μ M) and the fluorescence measured as a function of [PAP] in the presence or absence of $MgATP^{2-}$. Titration data were fitted to the usual binding equation, Equation 3, assuming a 1:1 stoichiometry, where F is the background-corrected fluorescence signal at 370 nm at each concentration of PAP, ΔF is the total change in fluorescence emission due to enzyme binding, C is the baseline-offset (i.e., fluorescence in the absence of PAP), $x = [A_3(2AP)A] = 1.95$ μ M, $y = [PAP]$ in μ M.

$$F = C + \frac{\Delta F \cdot (x + y + K_d) - \sqrt{(x + y + K_d)^2 - (4 \cdot x \cdot y)}}{2 \cdot x} \quad [3]$$

In the acrylamide quenching experiments, the background-corrected fluorescence emission of the probe was measured in the presence or absence of PAP and various nucleotides as a function of acrylamide concentration in the range of 0–0.4 M. The data were analyzed by non-linear regression analysis using the program gnuplot according to the Stern-Volmer equation: $F = F_0/(1 + K_{SV}[Q])$, where F is the fluorescence emission intensity at a given acrylamide concentration, Q ; F_0 is the maximum fluorescence intensity (in the absence of quencher); and K_{SV} is the Stern-Volmer constant. Because the addition of acrylamide to the protein and subsequent mixing of the sample resulted in precipitation, individual samples were prepared for each measurement at a given acrylamide concentration.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank M.C. Chang who grew the first crystals of crystal form 3 and Gretchen Meinke for her help throughout the project. We also give our appreciation to M. Forgac (Tufts University) for the use of his fluorescence spectrometer and to H. Robinson for data collection at NSLS beamline X29. Finally, we thank Claire Moore (Tufts) for critical reading of the manuscript. This work was supported by 5R01GM65972 to AB.

References

1. Hafe LA, Keller EB. The polyadenylate polymerases from yeast. *J Biol Chem* 1975;250:1838–46. [PubMed: 234467]
2. Lingner J, Radtke I, Wahle E, Keller W. Purification and characterization of poly(A) polymerase from *Saccharomyces cerevisiae*. *J Biol Chem* 1991;266:8741–6. [PubMed: 2026590]
3. Twu J, Bretthauer R. Properties of a polyriboadenylate polymerase isolated from yeast ribosomes. *Biochemistry* 1971;10:1576–1582. [PubMed: 5103999]
4. Wahle E. Purification and characterization of a mammalian polyadenylate polymerase involved in the 3' end processing of messenger RNA precursors. *J Biol Chem* 1991;266:3131–9. [PubMed: 1993684]
5. Barabino SM, Keller W. Last but not least: regulated poly(A) tail formation. *Cell* 1999;99:9–11. [PubMed: 10520989]
6. Chen J, Moore C. Separation of factors required for cleavage and polyadenylation of yeast pre-mRNA. *Mol Cell Biol* 1992;12:3470–81. [PubMed: 1352851]
7. Gross S, Moore C. Five subunits are required for reconstitution of the cleavage and polyadenylation activities of *Saccharomyces cerevisiae* cleavage factor I. *Proc Natl Acad Sci U S A* 2001;98:6080–5. [PubMed: 11344258]
8. Kessler MM, Zhao J, Moore CL. Purification of the *Saccharomyces cerevisiae* cleavage/polyadenylation factor I. Separation into two components that are required for both cleavage and polyadenylation of mRNA 3' ends. *J Biol Chem* 1996;271:27167–75. [PubMed: 8900210]
9. Minvielle-Sebastia L, Preker PJ, Wiederkehr T, Strahm Y, Keller W. The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. *Proc Natl Acad Sci U S A* 1997;94:7897–902. [PubMed: 9223284]
10. Zhao J, Kessler MM, Moore CL. Cleavage factor II of *Saccharomyces cerevisiae* contains homologues to subunits of the mammalian Cleavage/polyadenylation specificity factor and exhibits sequence-specific, ATP-dependent interaction with precursor RNA. *J Biol Chem* 1997;272:10831–8. [PubMed: 9099738]
11. Bienroth S, Wahle E, Suter-Crazzolara C, Keller W. Purification of the cleavage and polyadenylation factor involved in the 3'-processing of messenger RNA precursors. *J Biol Chem* 1991;266:19768–76. [PubMed: 1918081]

12. Christofori G, Keller W. 3' cleavage and polyadenylation of mRNA precursors in vitro requires a poly(A) polymerase, a cleavage factor, and a snRNP. *Cell* 1988;54:875–89. [PubMed: 2842067]
13. Christofori G, Keller W. Poly(A) polymerase purified from HeLa cell nuclear extract is required for both cleavage and polyadenylation of pre-mRNA in vitro. *Mol Cell Biol* 1989;9:193–203. [PubMed: 2538718]
14. Gilmartin G, Nevins J. An ordered pathway of assembly of components required for polyadenylation site recognition and processing. *Genes Dev* 1989;2:180–2190. [PubMed: 2628166]
15. Gilmartin G, Nevins J. Molecular analyses of two poly(A) site-processing factors that determine the recognition and efficiency of cleavage of the pre-mRNA. *Mol Cell Biol* 1991;11:2432–2438. [PubMed: 2017162]
16. Takagi Y, Ryner L, Manley J. Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation. *Cell* 1988;52:731–742. [PubMed: 2830992]
17. Takagi Y, Ryner L, Manley J. Four factors are required for 3'-end cleavage of pre-mRNA. *Genes Dev* 1989;3:1711–1724. [PubMed: 2558045]
18. Terns M, Jacob S. Role of poly(A) polymerase in the cleavage and polyadenylation of mRNA precursor. *Mol Cell Biol* 1989;9:1435–1444. [PubMed: 2566910]
19. Wahle E. Poly(A) tail length control is caused by termination of processive synthesis. *J Biol Chem* 1995;270:2800–8. [PubMed: 7852352]
20. Edmonds M. Polyadenylate polymerase. *Methods Enzymol* 1990;181:161–170. [PubMed: 2166211]
21. 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]
22. Sawaya MR, Pelletier H, Kumar A, Wilson SH, Kraut J. Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism. *Science* 1994;264:1930–5. [PubMed: 7516581]
23. Steitz TA, Smerdon SJ, Jager J, Joyce CM. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* 1994;266:2022–5. [PubMed: 7528445]
24. Aravind L, Koonin E. DNA polymerase β -like nucleotidyltransferase superfamily: Identification of three new families, classification and evolutionary history. *Nucleic Acids Res* 1999;27:1609–1618. [PubMed: 10075991]
25. Rittenhouse RC, Apostoluk WK, Miller JH, Straatsma TP. Characterization of the active site of DNA polymerase beta by molecular dynamics and quantum chemical calculation. *Proteins* 2003;53:667–82. [PubMed: 14579358]
26. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J. Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* 1994;264:1891–903. [PubMed: 7516580]
27. Burgers P, Eckstein F. A study of the mechanism of DNA polymerase I from *Escherichia coli* with diastereomeric phosphorothioate analogs of deoxyadenosine triphosphate. *J Biol Chem* 1979;254:6889–6893. [PubMed: 378995]
28. Sawaya MR, Prasad R, Wilson SH, Kraut J, Pelletier H. Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* 1997;36:11205–15. [PubMed: 9287163]
29. Steitz TA, Steitz JA. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A* 1993;90:6498–502. [PubMed: 8341661]
30. 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]
31. Zhelkovsky A, Helmling S, Moore C. Processivity of the *Saccharomyces cerevisiae* poly(A) polymerase requires interactions at the carboxyl-terminal RNA binding domain. *Mol Cell Biol* 1998;18:5942–51. [PubMed: 9742111]
32. Zhelkovsky AM, Kessler MM, Moore CL. Structure-function relationships in the *Saccharomyces cerevisiae* poly(A) polymerase. Identification of a novel RNA binding site and a domain that interacts with specificity factor(s). *J Biol Chem* 1995;270:26715–20. [PubMed: 7592899]

33. Martin G, Keller W, Doubie S. Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. *Embo J* 2000;19:4193–203. [PubMed: 10944102]
34. Helmling S, Zhelkovsky A, Moore CL. Fip1 regulates the activity of Poly(A) polymerase through multiple interactions. *Mol Cell Biol* 2001;21:2026–37. [PubMed: 11238938]
35. Kaufmann I, Martin G, Friedlein A, Langen H, Keller W. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *Embo J* 2004;23:616–26. [PubMed: 14749727]
36. Doubie S, Sawaya MR, Ellenberger T. An open and closed case for all polymerases. *Structure* 1999;7:R31–5. [PubMed: 10368292]
37. Joyce CM, Steitz TA. Function and structure relationships in DNA polymerases. *Annu Rev Biochem* 1994;63:777–822. [PubMed: 7526780]
38. Joyce CM, Steitz TA. Polymerase structures and function: variations on a theme? *J Bacteriol* 1995;177:6321–9. [PubMed: 7592405]
39. Martin G, Moglich A, Keller W, Doubie 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]
40. Kim SJ, Beard WA, Harvey J, Shock DD, Knutson JR, Wilson SH. Rapid segmental and subdomain motions of DNA polymerase beta. *J Biol Chem* 2003;278:5072–81. [PubMed: 12458221]
41. Dunlap CA, Tsai MD. Use of 2-aminopurine and tryptophan fluorescence as probes in kinetic analyses of DNA polymerase beta. *Biochemistry* 2002;41:11226–35. [PubMed: 12220188]
42. Frey MW, Sowers LC, Millar DP, Benkovic SJ. The nucleotide analog 2-aminopurine as a spectroscopic probe of nucleotide incorporation by the Klenow fragment of *Escherichia coli* polymerase I and bacteriophage T4 DNA polymerase. *Biochemistry* 1995;34:9185–92. [PubMed: 7619819]
43. Bose-Basu B, DeRose EF, Kirby TW, Mueller GA, Beard WA, Wilson SH, London RE. Dynamic characterization of a DNA repair enzyme: NMR studies of [methyl-¹³C]methionine-labeled DNA polymerase beta. *Biochemistry* 2004;43:8911–22. [PubMed: 15248749]
44. Beard WA, Wilson SH. Structure and mechanism of DNA polymerase Beta. *Chem Rev* 2006;106:361–82. [PubMed: 16464010]
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. Schneider TR. Objective comparison of protein structures: error-scaled difference distance matrices. *Acta Crystallogr D Biol Crystallogr* 2000;56:714–721. [PubMed: 10818348]
47. Schneider TR. Domain identification by iterative analysis of error-scaled difference distance matrices. *Acta Crystallogr D Biol Crystallogr* 2004;60:2269–75. [PubMed: 15572780]
48. Gerstein M, Lesk AM, Chothia C. Structural mechanisms for domain movements in proteins. *Biochemistry* 1994;33:6739–49. [PubMed: 8204609]
49. Zhelkovsky A, Helmling S, Bohm A, Moore C. Mutations in the middle domain of yeast poly(A) polymerase affect interactions with RNA but not ATP. *Rna* 2004;10:558–64. [PubMed: 15037764]
50. Bandwar RP, Patel SS. Peculiar 2-aminopurine fluorescence monitors the dynamics of open complex formation by bacteriophage T7 RNA polymerase. *J Biol Chem* 2001;276:14075–82. [PubMed: 11278877]
51. Rachofsky EL, Osman R, Ross JB. Probing structure and dynamics of DNA with 2-aminopurine: effects of local environment on fluorescence. *Biochemistry* 2001;40:946–56. [PubMed: 11170416]
52. Rai P, Cole TD, Thompson E, Millar DP, Linn S. Steady-state and time-resolved fluorescence studies indicate an unusual conformation of 2-aminopurine within ATAT and TATA duplex DNA sequences. *Nucleic Acids Res* 2003;31:2323–32. [PubMed: 12711677]
53. Ward DC, Reich E, Stryer L. Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J Biol Chem* 1969;244:1228–37. [PubMed: 5767305]
54. Gross S, Moore CL. Rna15 interaction with the A-rich yeast polyadenylation signal is an essential step in mRNA 3'-end formation. *Mol Cell Biol* 2001;21:8045–55. [PubMed: 11689695]

55. Colgan D, Manley J. Mechanism and regulation of mRNA polyadenylation. *Genes Dev* 1997;11:2755–2766. [PubMed: 9353246]
56. 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]
57. Martin G, Keller W. Sequence motifs that distinguish ATP(CTP):tRNA nucleotidyl transferases from eubacterial poly(A) polymerases. *Rna* 2004;10:899–906. [PubMed: 15146073]
58. Martin G, Jenö P, Keller W. Mapping of ATP binding regions in poly(A) polymerases by photoaffinity labeling and by mutational analysis identifies a domain conserved in many nucleotidyltransferases. *Protein Sci* 1999;8:2380–91. [PubMed: 10595540]
59. Kraynov VS, Showalter AK, Liu J, Zhong X, Tsai MD. DNA polymerase beta: contributions of template-positioning and dNTP triphosphate-binding residues to catalysis and fidelity. *Biochemistry* 2000;39:16008–15. [PubMed: 11123928]
60. Pedersen LC, Benning MM, Holden HM. Structural investigation of the antibiotic and ATP-binding sites in kanamycin nucleotidyltransferase. *Biochemistry* 1995;34:13305–11. [PubMed: 7577914]
61. Leslis AGW. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* 1992;26
62. Otwinowski ZMW. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 1997;276
63. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 1998;54:905–21. [PubMed: 9757107]
64. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 1994;50:760–3. [PubMed: 15299374]
65. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004;60:2126–32. [PubMed: 15572765]
66. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 1997;53:240–55. [PubMed: 15299926]
67. Bakhtina M, Lee S, Wang Y, Dunlap C, Lamarche B, Tsai MD. Use of viscogens, dNTPalphaS, and rhodium(III) as probes in stopped-flow experiments to obtain new evidence for the mechanism of catalysis by DNA polymerase beta. *Biochemistry* 2005;44:5177–87. [PubMed: 15794655]
68. Kraulis PJ. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystall* 1991;24:946–950.
69. DeLano, WL. The PyMOL User's Manual. San Carlos, CA: 2002.

Figure 1.

Summary difference distance matrix from ESCET. Error-weighted difference distance matrices were calculated for each of the ten unique pairs of different structures. The summary matrix shows the highest value at each matrix element among the ten matrices. Differences smaller than 4 sigma are colored grey. Those larger than 4 sigma are colored red (positive) or blue (negative), with the intensity of the color representing the magnitude of the shift. The most intense colors represent shifts of 15 sigma or larger. The large shifts near residue 270 are the result of a disordered loop at the surface of the protein.

Figure 2.

Superposition of five structures of yeast poly(A) polymerase. **(A)** The middle domains of the five structures were superimposed, so that movements about the two hinges can be observed. The molecules are colored as follows: 1A (blue), 1B (violet), 2A (cyan), 2B (turquoise), 3 (red). The palm domain, which opens and closes the central cleft, is on the left of the figure, and the two nucleotides at the incoming and 3'end positions in structure 1 are colored orange and gray, respectively. **(B)** Same as panel **(A)**, but rotated 90 degrees such that the C-terminal domain is in front and middle domain is at the bottom. The palm domain can just barely be seen in the fog behind the nucleotide. **(C)** Schematic view of the domain movements as seen from the top. These images were prepared using MOLSCRIPT⁶⁸.

A. **B.** **Figure 3.**

Model of the nucleotide-bound closed structure of PAP. This model shows the consequences of domain movements at the active site. **(A)** Molecules 1A and 3 are shown in blue and red, respectively, along with the nucleotides from 1A. The proposed nucleotide interaction with Lys215, Tyr224, and Asn226 are highlighted and the distances to the nucleotide in the open and closed model are noted with blue or black dashed lines, respectively. The catalytic (M1) and co-substrate (M2) metals are also depicted. This image was prepared using PyMOL⁶⁹. **(B)** Schematic representation of the same area as **(A)** depicting part of the hinge (Gly190) and the proposed substrate interactions (red arrows) upon further domain closure, as discussed in the text.

Figure 4.

Fluorescence emission spectra of the 2AP-labeled nucleic acids (**A**) $A_3(2AP)A$ and (**B**) $A_4(2AP)$. The concentrations of $A_3(2AP)A$ and $A_4(2AP)$ were $1.96 \mu M$ and $1.91 \mu M$, respectively. The color scheme in both panels is identical and refers to free nucleic acid alone (*black*) or in the presence of $105 \mu M$ PAP (*red*), PAP + AMP-CPP (*green*), PAP + CTP (*yellow*), or PAP + ATP (*blue*). The concentration of all nucleotides (in the Mg^{2+} -complex form) was $530 \mu M$, and the concentration of free Mg^{2+} was negligible. Spectra were taken immediately after sample preparation.

Figure 5.

Titration of $A_3(2AP)A$ with PAP in the absence (○) or presence (●) of $520 \mu M MgATP^{2-}$. The concentration of $A_3(2AP)A$ was $1.95 \mu M$; the concentration of PAP was in the range of 0.2 to $170 \mu M$. Samples were prepared individually. The background-corrected fluorescence intensity values (370 nm) are plotted and the solid line was generated from the nonlinear regression analysis of the data. The maximum fluorescence increase upon formation of the binary (○) or ternary complex (●) was 27% and 37% , respectively.

Figure 6.

Stern-Volmer plots of 2AP-labeled nucleic acids: **(A)** $A_3(2AP)A$ and **(B)** $A_4(2AP)$ in the presence or absence of PAP alone or PAP and various nucleotides. The color scheme and reagent concentrations were identical to those in Figure 4. The solid lines were generated from the parameters derived by non-linear regression analysis as described in Materials and Methods. The values for K_{SV} (M^{-1}) for $A_3(2AP)A$ are 3.06 ± 0.16 (free), 2.26 ± 0.10 (binary complex), 2.18 ± 0.09 (AMP-CPP ternary complex), 2.96 ± 0.09 (CTP ternary complex), 1.42 ± 0.02 (ATP ternary complex); for $A_4(2AP)$, 5.02 ± 0.26 (free), 3.30 ± 0.09 (binary complex), 3.42 ± 0.13 (AMP-CPP ternary complex), 3.16 ± 0.09 (CTP ternary complex), 1.67 ± 0.07 (ATP ternary complex).

Table 1

Crystallographic data and refinement statistics

	Crystal Form 2	Crystal Form 3	Form 1 (for comparison)
Space Group	P2 ₁ -2 ₁ -2 ₁	P2 ₁	P2 ₁ -2 ₁ -2 ₁
Unit cell dimensions a, b, c (Å)	82.22, 108.58, 132.72	56.16, 84.51, 71.32	73.8, 109.1, 238.5
α, β, γ (°)	90, 90, 90	90, 113.3, 90	90, 90, 90
Molecules per ASU	2	1	2
Solvent content	47.8 % ($V_M = 2.4 \text{ Å}^3/\text{Da}$)	50.2 % ($V_M = 2.5 \text{ Å}^3/\text{Da}$)	70.0% ($V_M = 4.1 \text{ Å}^3/\text{Da}$)
Resolution limit	2.8 Å	1.8 Å	2.6 Å
Observed (unique) reflections	323,403 (31,665)	394,549 (54,566)	
X-ray Source	NSLS beamline X25	NSLS beamline X29	
Multiplicity (last shell)	10.8 (6.2)	6.9 (5.7)	
Completeness (last shell)	99.3% (97.7%)	99.6% (96.9%)	
R_{merge} (last shell)	10.0% (30.6%)	5.7% (44.8)	
I/σ (last shell)	16.9 (5.2)	46.0 (4.3)	
Model Refinement			
R_{work} (R_{free})	19.50% (27.51%)	19.70% (23.55%)	
TLS groups	6	3	
Non-hydrogen atoms			
Protein	8351	4193	
Solvent	183	379	
Average B-factors (Å^2)			
Protein	26.7	30.4	
Solvent	13.0	38.4	
Ramachandran plot			
Most favored	848 (90.8%)	415 (92.0%)	
Outliers	5 (0.5%)	3 (0.7%)	
RMS deviations from ideal geometry			
Bond lengths	0.011 Å	0.017 Å	
Bond angles	1.318°	1.56°	
PDB deposition code	201P	2HHP	1FA0

* Each domain was defined to be a separate TLS group.
** B-factors were calculated using the program TLSANAL prior to calculation of the averages.

RMSD values for domain superimpositions and angular movements between Palm and C terminal domains. The boxes above the diagonal are values for the Palm domain movements relative to the middle. Values below the diagonal are for the C-terminal domain relative to the middle. Palm domain alignments involved residues 40-190 (53-203 in bovine PAP). C-terminal domain alignments among the yeast structures involved residues 353-427, 49-469 and 479-523. C-terminal alignments between the yeast and bovine structures involved 535-569 (bovine 366-382), 372-409 (bovine 384-421), 460-471 (bovine 433-444), 480-494 (bovine 455-469), and 505-516 (bovine 481-492).

Table 2

	1A	1B	2A	2B	3	1F5A
1A		4.77° 0.31Å	5.35° 0.55Å	4.02° 0.59Å	8.62° 0.41Å	6.33° 1.11Å
1B	1.38° 0.79Å		1.12° 0.49Å	2.72° 0.52Å	5.26° 0.39Å	10.33° 1.15Å
2A	2.96° 0.75Å	2.50° 0.51Å		2.38° 0.45Å	5.90° 0.49Å	10.81° 1.25Å
2B	2.68° 0.75Å	2.66° 0.79Å	5.11° 0.70Å		7.95° 0.48Å	10.38° 1.26Å
3	5.08° 0.59Å	5.22° 0.55Å	7.10° 0.53Å	4.33° 0.57Å		12.05° 1.10Å
1F5A	7.37° 1.11Å	8.90° 1.35Å	10.30° 1.22Å	5.52° 1.23Å	1.23° 6.40Å	