Protein Science (1999), *8*:2380–2391. Cambridge University Press. Printed in the USA.

Mapping of ATP binding regions in poly(A) polymerases by photoaffinity labeling and by mutational analysis identifies a domain conserved in many nucleotidyltransferases

GEORGES MARTIN,[2] PAUL JENÖ,[1] AND WALTER KELLER[1]

[1]Department of Cell Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland
[2]Department of Biochemistry, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

(RECEIVED July 19, 1999; ACCEPTED August 19, 1999)

Abstract

We have identified regions in poly(A) polymerases that interact with ATP. Conditions were established for efficient cross-linking of recombinant bovine and yeast poly(A) polymerases to 8-azido-ATP. Mn^{2+} strongly stimulated this reaction due to a 50-fold lower K_i for 8-azido-ATP in the presence of Mn^{2+}. Mutations of the highly conserved Asp residues 113, 115, and 167, critical for metal binding in the catalytic domain of bovine poly(A) polymerase, led to a strong reduction of cross-linking efficiency, and Mn^{2+} no longer stimulated the reaction. Sites of 8-azido-ATP cross-linking were mapped in different poly(A) polymerases by CNBr-cleavage and analysis of tryptic peptides by mass spectroscopy. The main cross-link in *Schizosaccharomyces pombe* poly(A) polymerase could be assigned to the peptide DLELSDNNLLK (amino acids 167–177). Database searches with sequences surrounding the cross-link site detected significant homologies to other nucleotidyltransferase families, suggesting a conservation of the nucleotide-binding fold among these families of enzymes. Mutations in the region of the "helical turn motif" (a domain binding the triphosphate moiety of the nucleotide) and in the suspected nucleotide-binding helix of bovine poly(A) polymerase impaired ATP binding and catalysis. The results indicate that ATP is bound in part by the helical turn motif and in part by a region that may be a structural analog to the fingers domain found in many polymerases.

Keywords: 8-azido-ATP; 3′-ends; helical turn motif; mRNA; nucleotidyltransferase; UV cross-linking

In eukaryotes, poly(A) polymerase (PAP) is part of a larger complex of *trans*-acting factors that processes the 3′-ends of newly transcribed messenger RNA precursors (pre-mRNA; Wahle & Keller, 1996; Colgan & Manley, 1997; Keller & Minvielle-Sebastia, 1997; Wahle & Rüegsegger, 1999). Many of the 3′-end processing factors identified in mammals have counterparts in *Saccharomyces cerevisiae* (Manley & Takagaki, 1996; Keller & Minvielle-Sebastia, 1997).

Poly(A) polymerases have been cloned from *S. cerevisiae* (Lingner et al., 1991), bovine (Raabe et al., 1991; Wahle et al., 1991), human (Thuresson et al., 1994), *Xenopus laevis* (Ballantyne et al., 1995; Gebauer & Richter, 1995), *Drosophila melanogaster* (F. Juge & M. Simonelig, pers. obs.), and vaccinia virus (Gershon et al., 1991). PAP genes from two other yeasts were also cloned: the *Schizosaccharomyces pombe pla1* gene (Ohnacker et al., 1996), and the *Candida albicans* PAP homologue (Ishii et al., 1997). PAP homologues from *Caenorhabditis elegans* (Gardner, 1995) and rice (Yamamoto & Sasaki, 1997) are found in the sequence databases.

Bacterial poly(A) polymerases (Cao & Sarkar, 1992), which also contain a similar catalytic fold as the eukaryotic and vaccinia PAPs (Martin & Keller, 1996), were found to be associated with polysomes and are involved in mRNA decay (Ingle & Kushner, 1996). The enzyme cca:tRNA nucleotidyltransferase (ccaT), a close relative of bacterial PAP, is involved in maturation of tRNA, and there are no characteristic sequence features by which these two enzyme classes could be distinguished (Yue et al., 1996).

Based on sequence homology and mutational analysis of bovine PAP (Holm & Sander, 1995; Martin & Keller, 1996), poly(A) polymerases were found to be members of a superfamily of nucleotidyltransferases that are characterized by a sequence motif

Reprint requests to: Walter Keller, Department of Cell Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, Basel, CH-4056 Switzerland; e-mail: walter.keller2@unibas.ch.

Abbreviations: ccaT, cca:tRNA nucleotidyltransferase; DTT, 1,4-dithiothreitol; EDTA, ethylenedinitrilo tetraacetic acid; HTM, helical turn motif; KanNt, kanamycin nucleotidyltransferase; PAP, poly(A) polymerase; PCR, polymerase chain reaction; Pol β, DNA polymerase β; Pol I, DNA polymerase I; pre-mRNA, messenger RNA precursors; PVDF, polyvinylidene difluoride; TCA; trichloroacetic acid; TdT, terminal deoxynucleotidyl transferase; TFA, trifluoracetic acid.

with the consensus G[G/S]$(x)_{9-13}$Dx[DE] (x stands for any nucleotide). A helical loop structure is formed by the first four amino acids of the sequence motif starting with the invariant G[G/S]. This helical turn motif (HTM) or mononucleotide binding domain (Sawaya et al., 1997) was found in the DNA polymerase β (Pol β) and kanamycin nucleotidyltransferase (KanNt) structure; therefore, all members of this nucleotidyltransferase superfamily (hereafter called HTM-β superfamily to distinguish from other superfamilies of nucleotidyltransferases or polymerases) were predicted to contain this structure (Holm & Sander, 1995; Martin & Keller, 1996). Homologies that conform to the HTM-β motif were also found in the bacterial PAPs and ccaTs as well as in the vaccinia PAP sequences (Martin & Keller, 1996). Further members of the HTM-β-superfamily were detected by homology searches: terminal deoxynucleotidyl transferase (TdT), 2′-5′ oligo(A) synthetase and the antibiotic resistance factor streptomycin 3″-adenylyltransferase as well as protein nucleotidyltransferases like protein-P_{II} uridylyltransferase and glutamine synthetase adenylyltransferase (Holm & Sander, 1995). The HTM-β superfamily of nucleotidyltransferases was further subdivided into three subclasses, depending on the type of sequence homology within the characteristic helical turn motif (Yue et al., 1996). Three new members of the superfamily were found recently: a minimal nucleotidyltransferase, the TRF4/5 proteins, and a family involved in signal transduction (Aravind & Koonin, 1999).

So far, no crystal structure of any poly(A) polymerase has been reported. On the other hand, DNA polymerase β from humans and rats was crystallized and its structure solved (Davies et al., 1994; Pelletier et al., 1994; Sawaya et al., 1994). The crystal structure of kanamycin nucleotidyl transferase has also been solved (Sakon et al., 1993; Pedersen et al., 1995), and the similarity of its catalytic fold to Pol β was confirmed. This enzyme modifies kanamycin by attaching an AMP molecule to the 4′-OH group of the antibiotic.

The structure of most polymerases was found to have the shape of a half-open human hand, forming a cleft between the fingers and the thumb, and structural subdomains were called palm, fingers, and thumb domain (Ollis et al., 1985). Moreover, comparison of the nucleotide binding pockets of DNA polymerase I (Pol I) and Pol β revealed conserved features in their active site geometry. In both enzymes, the triphosphate wraps around the metal at site A and at the same time the base pairs with the template. Pol β (Sawaya et al., 1997) and polymerases of the Pol I type from *Thermus aquaticus* (Li et al., 1998) and T7 DNA polymerase (Doublié et al., 1998) were crystallized as a ternary complex with a DNA primer, a template strand, and a dideoxynucleoside triphosphate. The main difference between Pol β and Pol I-type polymerases consists in the source of interaction with the phosphoryl oxygens. In Pol I, residues located on the N-terminal half of helix O form hydrogen bonds with the α-, β-, and γ-phosphoryl oxygens. In Pol β, this is achieved by residues in the helical turn motif. Both the crystal structure of Pol β and KanNt (Davies et al., 1994; Pelletier et al., 1994; Pedersen et al., 1995) have shown that residues in this motif are responsible for binding the triphosphate moiety of the incoming nucleotide. In T7 DNA polymerase, the ribose of the nucleotide is positioned by two residues, Glu480 and Tyr526 (Doublié et al., 1998). Tyr530, although less important, is another residue on the O-helix that helps to position the nucleotide. In Pol β, residues Tyr271, Phe272, and Asn279 on helices M and N interact with the ribose and the base of the nucleotide (Pelletier et al., 1994), and the two helices of the fingers domain were found to participate in an induced fit mechanism (Sawaya et al., 1997).

Another approach to study nucleotide binding—in particular, when a structure is not available—is photoaffinity labeling (Potter & Haley, 1983). In a number of studies, the aryl-azide photo probes azido-ATP (N_3ATP), N_3dATP, or N_3dGTP were used with the α- or γ-position of the triphosphate labeled with ^{32}P. Sites of interaction with a photoprobe were investigated in the Klenow fragment of DNA polymerase I and [γ-^{32}P]-8-N_3dATP (Rush & Konigsberg, 1990) and terminal deoxynucleotidyl transferase and [γ-^{32}P]-8-N_3dATP (Evans et al., 1989). DNA polymerase β and both [γ-^{32}P]-8-N_3ATP and [γ-^{32}P]-5-N_3dUTP (Srivastava et al., 1996) and HIV-1 reverse transcriptase and [methyl-^{3}H]dTTP (Cheng et al., 1993).

In this paper, we used the photolabile ATP analog 8-azido-ATP to confirm and extend previous results (Martin & Keller, 1996), which suggested the involvement of the aspartate residues 113, 115, and 167 of bovine PAP in the coordination of two catalytic metal ions. We mapped the region of the cross-link in poly(A) polymerases after cleavage with cyanogen bromide and trypsin. By homology search, we identified a group of proteins that are related to the region of the cross-link in PAP. We also generated a number of point mutations in this region in PAP to identify residues that interact with the incoming ATP molecule. At the same time we extended the mutagenesis to a region in PAP that is likely to participate in binding the triphosphate moiety of the nucleotide.

Results

Specificity of 8-azido-ATP interaction with a nucleotide binding domain of PAP

To investigate the suitability of N_3ATP as a substrate for poly(A) polymerase, we tested the incorporation of unlabeled N_3ATP into the growing poly(A) chain. The bulky azido group (−N=N=N) of the photo probe and, after photoactivation, the smaller but highly reactive nitrene group can be expected to interfere with binding by steric hindrance. In the experiment shown in Figure 1, one or two molecules of N_3ATP were incorporated (lane 2). Higher amounts of PAP did not lead to longer products but caused some degradation (lane 3). Ultraviolet (UV) irradiation was slightly stimulatory, probably by releasing two nitrogens from the azido group (lane 4). The addition of DTT had an even stronger effect possibly by quenching most of the azido groups (lane 5). With the same concentration of regular ATP, PAP elongated the poly(A) chain by about 60–70 residues (lane 6). When N_3ATP was mixed with an equal amount of ATP, PAP was not noticeably inhibited (lane 7). N_3ATP is, therefore, a weak competitive inhibitor for PAP. Although the nucleotide analog inhibits the reaction, it binds to the enzyme reversibly. We have measured an inhibitor constant $K_{i(N_3ATP)}$ of 2.6 mM in the presence of Mg^{2+} and 0.05 mM with Mn^{2+} (results not shown).

To show that N_3ATP cross-linked to genuine ATP-binding regions and not to nonspecific sites on the protein, competition experiments were carried out with various nucleoside triphosphates (Fig. 2A). A 100-fold excess of the ribonucleoside triphosphates ATP, CTP, GTP, and UTP and the deoxynucleoside triphosphates dATP, dCTP, dGTP, and dTTP were added to the reaction (lanes 2–9). Competition with all the nucleotides tested was between 80 and 90%. Competition with ATP and dATP was not significantly different from competition with the other nucleotides. Strong inhibition by ribo- and deoxyribonucleoside triphosphates of the activity of purified poly(A) polymerases has previously been observed (E. Wahle, pers. obs.; Hadidi & Sethi, 1976; Ed-

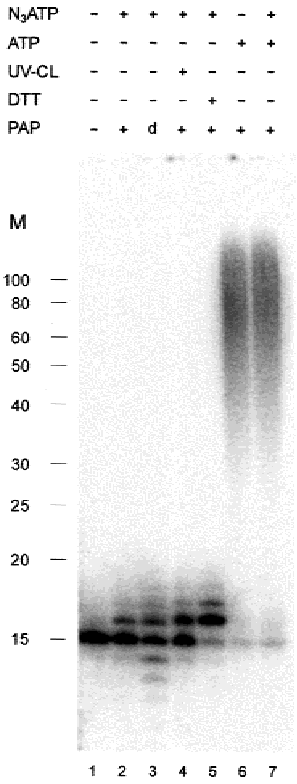

Fig. 1. Specificity and efficiency of UV cross-linking 8-azido-ATP with different poly(A) polymerases. Incorporation of N_3ATP by PAP into an $(A)_{15}$ primer was measured. The standard reaction contained, when indicated by a +, 0.02 mM N_3ATP, 0.02 mM ATP, 10 mM DTT, and 1.7 pmol bovine PAP-513 (except in lane 3 where "d" denotes addition of 17 pmol of PAP-513). The concentration of the primer $(A)_{15}$ was 1 μM. UV-CL stands for 1 min of UV cross-linking before PAP was added. One-quarter of the reaction was separated on a 20% sequencing gel and exposed to a PhosphorImager screen.

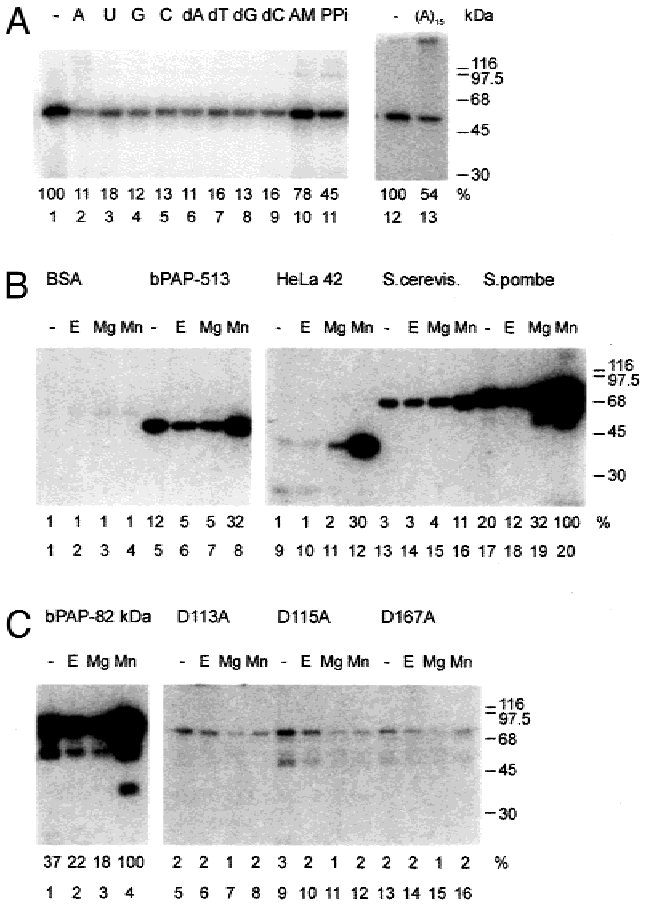

Fig. 2. Specificity test and quantitation of UV cross-linking of N_3ATP with poly(A) polymerases. **A:** Competition of UV cross-linking was measured in reactions that contained in cross-link buffer in a total volume of 10 μL:10 pmol protein bPAP-513, 1.25 μM [γ-^{32}P]N_3ATP and 125 μM competitor nucleotides or $(A)_{15}$ (lane 13). Reactions were separated on 10% SDS-PAGE. **B:** Stimulation of UV cross-linking by Mg^{2+} and Mn^{2+}. Reactions contained in 10 μL: cross-link buffer, bovine serum albumin (lanes 1–4, designated BSA at the top of the panel), 513 N-terminal fragment of bovine PAP (lanes 5–8, bPAP-513) HeLa 42 kDa PAP (lanes 9–12, HeLa42), *S. cerevisiae* PAP (lanes 13–16, *S. cerevis*), and *S. pombe* PAP (lanes 17–20, *S. pombe*). Additions to reactions consisted of 10 mM EDTA (E in lanes 2, 6, 10, 14, and 18), 4 mM $MgCl_2$ (Mg in lanes 3, 7, 11, 15, and 19) and 0.5 mM $MnCl_2$ (Mn in lanes 4, 8, 12, 16, and 20). **C:** Stimulation of UV cross-linking of catalytic aspartate mutants. Reactions were performed as those in **B**. PAP was added as the 82 kDa bovine PAP (lanes 1–4), D113A mutant (lanes 5–8), D115A mutant (lanes 9–12), and D167A mutant (lanes 13–16). Reactions in **B** and **C** were separated on 10% SDS-PAGE gels. Gels were exposed to PhosphorImager screen and signals were quantified with IPLabGel software (Signal Analytics Corporation).

monds, 1982). Our results suggest that binding of the nucleoside triphosphates takes place predominantly between the phosphate moieties of the nucleotide and PAP. The contribution to binding by the base before the transition state, at the stage when competition of cross-linking was tested, is probably minor, and it is likely that the selection of ATP during incorporation is accomplished by a conformational change of the enzyme at the transition state.

AMP and pyrophosphate (PP_i) were also included as competitors (lanes 10 and 11). AMP competed weakly under these conditions, confirming an important role for the triphosphate moiety of the nucleotide for binding specificity. Pyrophosphate, on the other hand, was somewhat more effective in competing for the N_3ATP molecule, probably by occupying the site for binding of the triphosphate moiety of ATP.

In addition, the effect of an excess of the RNA $(A)_{15}$ was tested (Fig. 2A, right panel). $(A)_{15}$ inhibited photo-labeling less strongly than the nonspecific nucleotides. We often observed a smear of label between the gel slot and the band representing the protein-N_3ATP cross-link, which suggests that some of the primer $(A)_{15}$ was also cross-linked to PAP and slowed down migration of the protein in the gel.

We concluded that UV cross-linking with N_3ATP is a valid means of assaying the ATP-binding domain of poly(A) polymerase.

UV cross-linking of yeast and bovine poly(A) polymerases with N_3ATP

Recombinant poly(A) polymerases of calf, *S. cerevisiae* and *S. pombe*, were tested for UV cross-linking with the photoprobe N_3ATP under reaction conditions that either included the metal chelator EDTA or one of the divalent metals Mg^{2+} or Mn^{2+}. The 42 kDa HeLa PAP, an inactive form of poly(A) polymerase (Wahle et al., 1991), was included in the experiment because we anticipated this protein to contain a complete catalytic core but to lack the RNA binding domain. This short form of PAP is inactive in an in vitro assay (Martin & Keller, 1996).

Addition of EDTA sometimes caused a slight reduction of the cross-link (Fig. 2B), probably by removal of metal ions bound to the catalytic center of the polymerases (relative amounts of cross-links are indicated below the autoradiographs). When Mg^{2+} was included in the photolabeling reaction, a moderate stimulation of the cross-linking efficiency was observed with most of the PAPs, but only when compared to the reactions that contained EDTA. A much stronger stimulation of cross-linking was seen when Mn^{2+} was present. The reason for this phenomenon seems to be the much lower K_i for N_3ATP in the presence of Mn^{2+}, as mentioned above. Differences of the stimulation between the various PAPs were reproducible and are characteristic for each enzyme. *S. pombe* PAP showed the highest efficiency of cross-linking, which was up to 10 times higher than that obtained with the other PAPs. The 42 kDa HeLa PAP could be cross-linked with approximately the same efficiency as the bovine or *S. cerevisiae* PAPs, and it showed the same response to addition of Mn^{2+}. This confirms that an intact ATP binding domain and likely the entire catalytic fold is present in this short form of PAP. Strikingly, almost no cross-linking occurred with the 42 kDa HeLa PAP in the absence of divalent cations or in the presence of EDTA. No signal was detected with BSA as a control in cross-linking reactions.

UV cross-linking of active site mutants with N_3ATP

We tested mutant bovine poly(A) polymerases with aspartate residues 113, 115, and 167 mutated to alanine. Mutations in these residues severely affect catalysis (Martin & Keller, 1996). Aspartates are found in analogous positions as in the Pol β catalytic center where structural analysis demonstrated that these residues interact with metal ions (Pelletier et al., 1994). For these experiments, the aspartate mutants of the 82 kDa form of bovine poly(A) polymerase were used together with a 82 kDa wild-type control (Fig. 2C, left panel). The mutants showed no stimulation of N_3ATP cross-linking when Mg^{2+} or Mn^{2+} was included in the reaction (Fig. 2C, right panel). In fact, addition of either of the two metals caused an inhibition of the weak cross-link. Interestingly, the presence of either Mn^{2+} or Mg^{2+} to the reaction resulted in a more reduced cross-link compared to reactions without any addition. ATP in solution together with Mn^{2+} or Mg^{2+} is expected to form a Me^{2+}–ATP complex (Sawaya et al., 1997), and apparently this complex binds less well than N_3ATP without metal ion in these mutants.

These results also confirm our assumption that amino acid Asp167 participates in the coordination of the two metals as found in the Pol β structure. If Asp167 is needed for metal coordination, its mutation to alanine, which also abolishes catalysis, may also affect cross-linking of N_3ATP.

CNBr mapping of PAPs cross-linked with N_3ATP

To map the location of the cross-link on the polypeptide chain, recombinant poly(A) polymerases from calf, *S. cerevisiae* and *S. pombe*, were treated after photolabeling with cyanogen bromide (CNBr; Nikodem & Fresco, 1997), an agent that cleaves proteins at the carboxyl side of methionine residues (Fig. 3A,B). Peptide fragments were separated on Tricine-SDS-PAGE gels (Schägger & von Jagow, 1987) and blotted to PVDF membranes. It has been shown that small molecular weight protein fragments are lost from gels during fixation but blot well to membranes (Plough et al., 1989). We found that with each PAP one fragment was reproducibly labeled strongly compared to several weaker bands, although

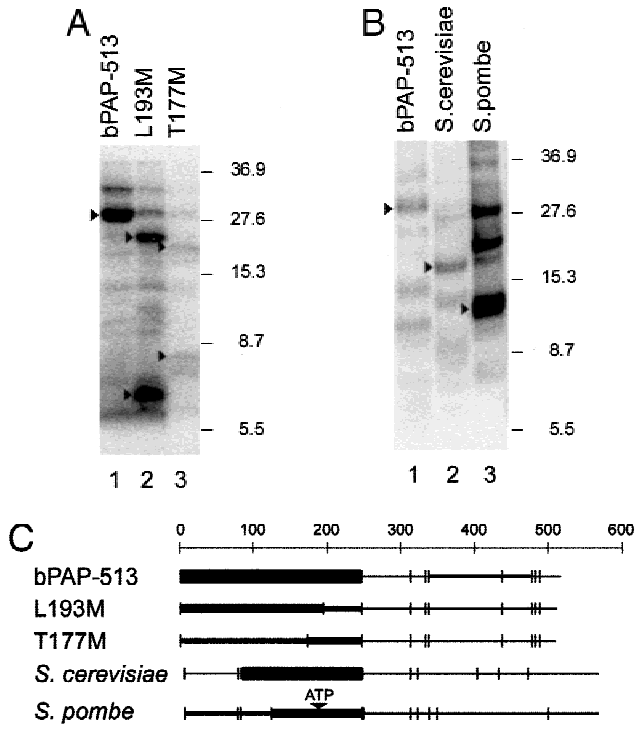

Fig. 3. CNBr mapping of poly(A) polymerases cross-linked to N_3ATP. Poly(A) polymerase (20 μg) was cross-linked with N_3ATP and treated with CNBr. The fragments were separated on a 16.5% tricine gel and blotted to a PVDF membrane that was then exposed to a PhosphorImager screen. **A:** Separation of bovine PAP C-terminal deletion (bPAP-513) and two mutants with residues Leu193 and Thr177 mutated to methionines. Black arrowheads point to the major cross-links (or two fragments if about equally labeled). **B:** Comparison of labeled CNBr fragments of bPAP-513 with fragments of *S. cerevisiae* and *S. pombe*. Black arrowheads point to fragments with major cross-links. Gels were exposed to PhosphorImager screen and signals were quantified with IPLabGel software. **C:** Map of CNBr digests in **A** and **B**. Horizontal lines symbolize linear polypeptide chains of respective PAPs. Lengths indicate sizes of the proteins in amino acids according to the scale at the top. Thicker horizontal bars indicate relative intensity of the label in CNBr fragments as an average of several experiments. Vertical short lines signify CNBr cleavage-sites (methionines).

the difference in intensity to the minor bands was not always large (Fig. 3). The fragment with the major cross-link usually contained between 40 and 60% of the label of all fragments. The major cross-link in bovine and *S. cerevisiae* PAP was found in CNBr fragments of 28.5 and 18.1 kDa, respectively, which represent the largest possible CNBr fragments in the respective poly(A) polymerases (Fig. 3B). In *S. pombe* PAP, the labeled fragment corresponds to the 12.9 kDa fragment, which extends from amino acids 122 to 237. Fragments migrating slower on the gel represent products of partial cleavage. Weaker bands could not be assigned unambiguously, and they could also be the products of partial digestion.

Initially, we were mainly interested in the site of the N_3ATP cross-link in bovine PAP. Unfortunately, peptides from bovine PAP could not be investigated by mass spectrum analysis (see below) because this protein cross-linked less efficiently with N_3ATP than *S. pombe* PAP. For this reason, another method was used to finemap bovine PAP, whereby mutants with artificial methionines were generated and fragmented with CNBr. In results obtained after cross-linking the two methionine mutants L193M and T177M,

CNBr was found to cleave the large fragment that spans from residues 1 to 251 into two subfragments (Fig. 3A, lanes 2 and 3). The label was found at about equal intensity in the two fragments, indicating that more than one cross-link must exist in this part of the polypeptide. When the experiment was repeated with the mutant L193M, 87% of the label shared by the two fragments was found in the large fragment. The reason for these variations must be small differences in the conditions during cross-linking.

Figure 3C shows a graphical summary of the results from the CNBr mapping experiments. Labeling intensities are indicated by the thickness of the lines symbolizing the protein fragments. They are an average of all experiments and do not only represent those shown in Figure 3A.

Mass spectral analysis of N_3ATP-modified tryptic peptides

In an attempt to identify peptides that carry cross-linked azido-ATP, photo labeled *S. pombe* PAP was digested with trypsin. Because UV cross-linking with *S. pombe* PAP was more efficient than with *S. cerevisiae* or bovine PAP, the experiments described below were done with the *S. pombe* enzyme exclusively. Incorporation of label was between 10 and 20%, as measured after TCA precipitation of a small aliquot after cross-linking. A large part of the radioactivity (60–95%) was found in the flowthrough when the digest was separated by reverse-phase chromatography. The nature of the radioactive material in the flowthrough is not known, but it has been shown that peptides cross-linked with aryl azides were unstable when fractionated by reversed-phase high-performance liquid chromatography in a TFA containing buffer system (Salvucci et al., 1992; Cheng et al., 1993). We believe that at least part of the released label represents free phosphate originating from the cross-linked nucleotide because some of the radioactivity applied to TLC plates comigrated with radioactive phosphate (data not shown). A labile glycosidic bond in the azide-nucleotide has also been suspected to cause loss of label (Cheng et al., 1993). Due to the high losses of radioactivity, the large amounts of peptides used in these experiments required crude prefractionation on a RESOURCE-RPC reversed-phase column (Fig. 4A). In the experiment shown, several peaks eluted, with major peaks at about 37 and 70% solvent B, respectively. The reason for this is most likely an incomplete digestion of the protein, and when the digest was repeated, one major peak at 35 to 40% solvent B was obtained. Fractions containing labeled material (indicated by a frame) were then applied to a VYDAC C_{18} column, where resolution was better, and interestingly, no additional loss of label occurred (Fig. 4B). When the fractions containing labeled material (from 29 to 33 min) were analyzed by mass spectrometry, a complex spectrum was obtained, indicating a mixture of different peptides (Fig. 4C). Close inspection of the spectrum revealed a pair of signals with the mass to charge ratio of 898.7 and 1,795.6 Da. This pair of signals corresponds to a doubly and singly charged peptide with a mass of 1,794.61 Da. Comparison of the calculated masses of each of the tryptic fragments of PAP with the measured mass of 1,795.6 Da showed that the peptide DLELSDNNLLK fits exactly the observed mass shift caused by the attachment of the nitrene–ATP moiety. In a second experiment, the same peptide–nitrene ATP adduct was identified as a series of signals corresponding to the triply (599.2 Da), doubly (898.6 Da), and the singly (1,795.5 Da) charged ion (not shown). Essentially, all the remaining signals could be assigned to tryptic peptides without attached nitrene–ATP. Attempts to locate the site of modification within the peptide more precisely were not successful because the cross-linking product was too unstable.

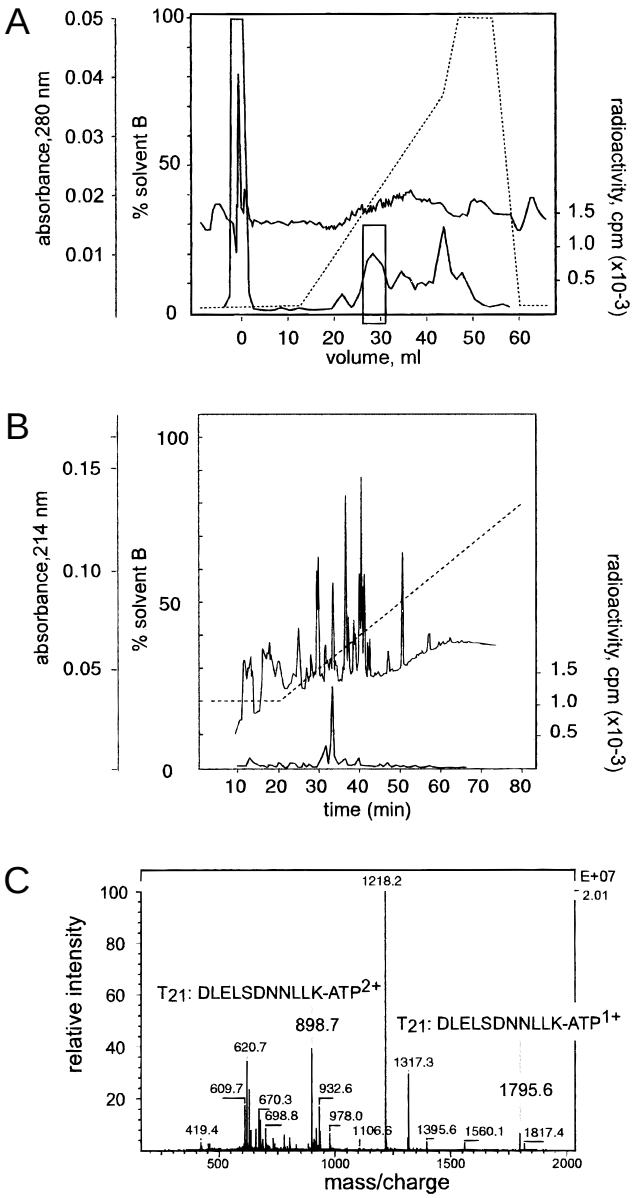

Fig. 4. Reversed-phase chromatography and mass spectral analysis of *S. pombe* PAP cross-linked to N_3ATP. Results from the same experiment are depicted in **A** to **C**. **A:** Chromatogram of a prefractionation of 100 μg of PAP peptides on a 1 mL RESOURCE-RPC column. The dotted line represents the concentration of solvent B (80% CH_3CN in 0.1% TFA). The lower curve is the profile of measured radioactivity in the fractions, and the profile in the center is UV absorbance. Fractions 26 to 31 (indicated by a frame) were pooled and rerun on a 2.1 × 250 mm VYDAC column. **B:** Chromatographic conditions were as in **A**. Fractions of 50 μL were collected. **C:** Mass spectrum of fraction 33 (**B**) showing the T_{21}-ATP–modified peptide as the doubly (898.7 Da) and singly (1,795.6 Da) charged ion.

Multiple sequence alignment of the putative nucleotide binding domains

With the results from CNBr cleavage and mass spectra analysis, we suspected that the region around amino acid 200 in poly(A)

polymerases has a role in nucleotide binding. We then performed database searches with the PSI-BLAST (Altschul et al., 1997) program at the NCBI server (website:www.ncbi.nlm.nih.gov). In the first search, we used the bovine PAP protein sequence from residues 94 to 245 as query, which includes both the nucleotidyltransferase signature G[G/S](x)$_{9\text{-}13}$Dx[D/E] and the suspected PAP ATP-binding motif. Positives were included in further iterations only if they contained significant homologies in the prospective nucleotide binding region. We hypothesized this domain to extend some 60 to 80 amino acids downstream of the single Asp motif. The resulting positives were members of the family of topoisomerase-related TRF4/5-type proteins involved in chromatin condensation (Castano et al., 1996). In the first iteration, the TRF4/5 proteins were now found within the $e = 0.001$ significance threshold, and the highest score was reached by TRF4 of *S. cerevisiae* with an e-value of 4e-18. Significant hits in the second iterative PSI-BLAST search were a large number of the members of the family of the streptomycin 3″-adenylyltransferases.

For a second PSI-BLAST search, we employed the sequence of amino acids 220 to 380 of the *S. cerevisiae* protein TRF-5 (Swiss-Prot ID: trf5_ yeast). The first iteration returned homologies with several 2′-5′ oligo(A) synthetases with a best e-value of 4.7. This family of proteins then appeared well within the threshold in the second iteration with an e-value of 6e-09. In the third iteration, several members of the archaeal cca:tRNA nucleotidyl transferases appeared above the threshold with an e-value of 6e-21 for the highest scoring member.

The multiple sequence alignment in Figure 5 illustrates conservation of sequence blocks between different groups of enzymes and conserved residues are highlighted by different colored backgrounds (see caption for details). The group of Pol β is depicted for comparison, and very little homology is found between this segment, which represents the fingers domain in Pol β and any other protein. For the fingers and thumb domains, we are using the nomenclature of Steitz (Steitz et al., 1994). One common feature between Pol β and the TRF4/5 proteins could be the presence of two helices of about the same size. For comparison, we also included the sequence of the fingers domain of several Pol I proteins at the lower part of the alignment. A significant homology was not detected by the PSI-BLAST database searches, but surprisingly, we found a conservation of some key Pol I residues between the archaeal ccaTs and the Pol I sequences. The invariant His at position 38 located in the T7 DNA polymerase structure in the region of helix N was found to bind to one of the β-phosphoryl oxygens, and the conserved Tyr-Gly motif (position 70/71) is important to form a pocket for the incoming nucleotide on helix O in Pol I polymerases (Doublié et al., 1998).

Results of a structure prediction at the PHD server at the EMBL (Rost & Sander, 1993) support the hypothesis of a structural homology between these classes of proteins. The prediction program calculated a strong tendency for helical structures in the region corresponding to helices N and O of Pol I in each separate family of transferases (Fig. 5, >phd line below groups). This strongly argues in favor of a conservation of these two helices between the different protein families.

Activity and steady-state kinetics of mutants in the helical turn motif

To test interactions of residues in the helical turn motif, point mutations were generated between residues 100 and 105 in the 58 kDa bovine PAP C-terminal deletion construct (Martin & Keller, 1996). Ser180, Arg183, and Gly189 form hydrogen bonds with β- and γ-phosphate oxygens of the incoming nucleotide in Pol β (Pelletier et al., 1994; Sawaya et al., 1997). Mutants of analogous residues in bovine PAP were constructed, and the recombinant proteins were expressed in *Escherichia coli*. Measurement of kinetic parameters indicate that in several of the mutant proteins both $K_{M(\mathrm{ATP})}$ and k_{cat} are affected (Table 1). The rate constant k_{cat} of mutants S102A, R104A, and L105A is strongly increased compared to wild-type PAP. A reason for this could be that because these residues were all replaced by alanine, an obstacle was removed that could lead to either a more efficient ATP influx or a faster release of PP_i after formation of the phosphodiester bonds.

Furthermore, measurement of the inhibition of nonspecific PAP activity by pyrophosphate (Table 2) showed that mutations F100D and S102A rendered the polymerase sensitive to pyrophosphate. About threefold less pyrophosphate was necessary to reduce the reaction to half compared to wild-type PAP. This is consistent with the assumption of an interaction of Ser102 in the helical turn motif with the phosphate moiety of the ATP molecule. Phe100 is very likely the analog of Tyr37 in KanNt, a residue that acts as a supporting element for the ribose (Pedersen et al., 1995). Addition of pyrophosphate to the F100D mutant could destabilize NTP binding further, when the presence of Asp at position 100 is unfavorable compared to Phe for interactions with the nucleotide.

Mutagenesis of residues in the suspected fingers domain and kinetic constants of mutants

A number of site-directed point mutants in the region between amino acids 172 and 243 of the bovine PAP 58 kDa deletion mutant (bPAP-513) were generated in vitro. The rationale for mutating Phe220 and Lys228 was the conservation of these residues in Pol β and TdT (Fig. 5). Steady-state kinetics showed that mutations F220R, F220S, and K228A caused an increase in $K_{M(\mathrm{ATP})}$ (Table 1). The increase was moderate, between four- and fivefold compared to wild-type. In a multiple-site mutant (NIDNF-216-SIANS), a 15-fold higher $K_{M(\mathrm{ATP})}$ was measured compared to wild-type. The conservative mutation F220Y caused no effect on $K_{M(\mathrm{ATP})}$, suggesting a preference for an aromatic residue at this position. In addition, the mutant K228A reduced k_{cat} about threefold, implying a role for this amino acid in nucleotide recognition as well as in catalyzing the polymerization reaction. Amino acids Arg225 and Lys228 in PAP were mutated to alanines based on their homology to residues shown to form contacts with the incoming nucleotide in DNA polymerase I type enzymes (Kiefer et al., 1997; Doublié et al., 1998; Li et al., 1998). The mutations W230A, Y237A, and F243A were made because at the the C-terminal end of helix O in DNA polymerases I, aromatic residues were found to play an important role in discriminating between 3′-OH and 3′-deoxynucleotides and to form a snug fit with the ribose and base of the incoming nucleotide (Joyce & Steitz, 1994; Tabor & Richardson, 1995; Astatke et al., 1998). All three mutations in PAP caused a three to five times higher $K_{M(\mathrm{ATP})}$ compared to wild-type, suggesting involvement in nucleotide binding.

A test of the inhibitory effect of pyrophosphate on PAP activity in these mutants showed very little or no difference compared to wild-type (Table 2). This supports the idea of a separate function of amino acids in the two domains investigated. The helical turn motif is thought to bind to the phosphoryl oxygens and the fingers domain to contact the nucleotide base.

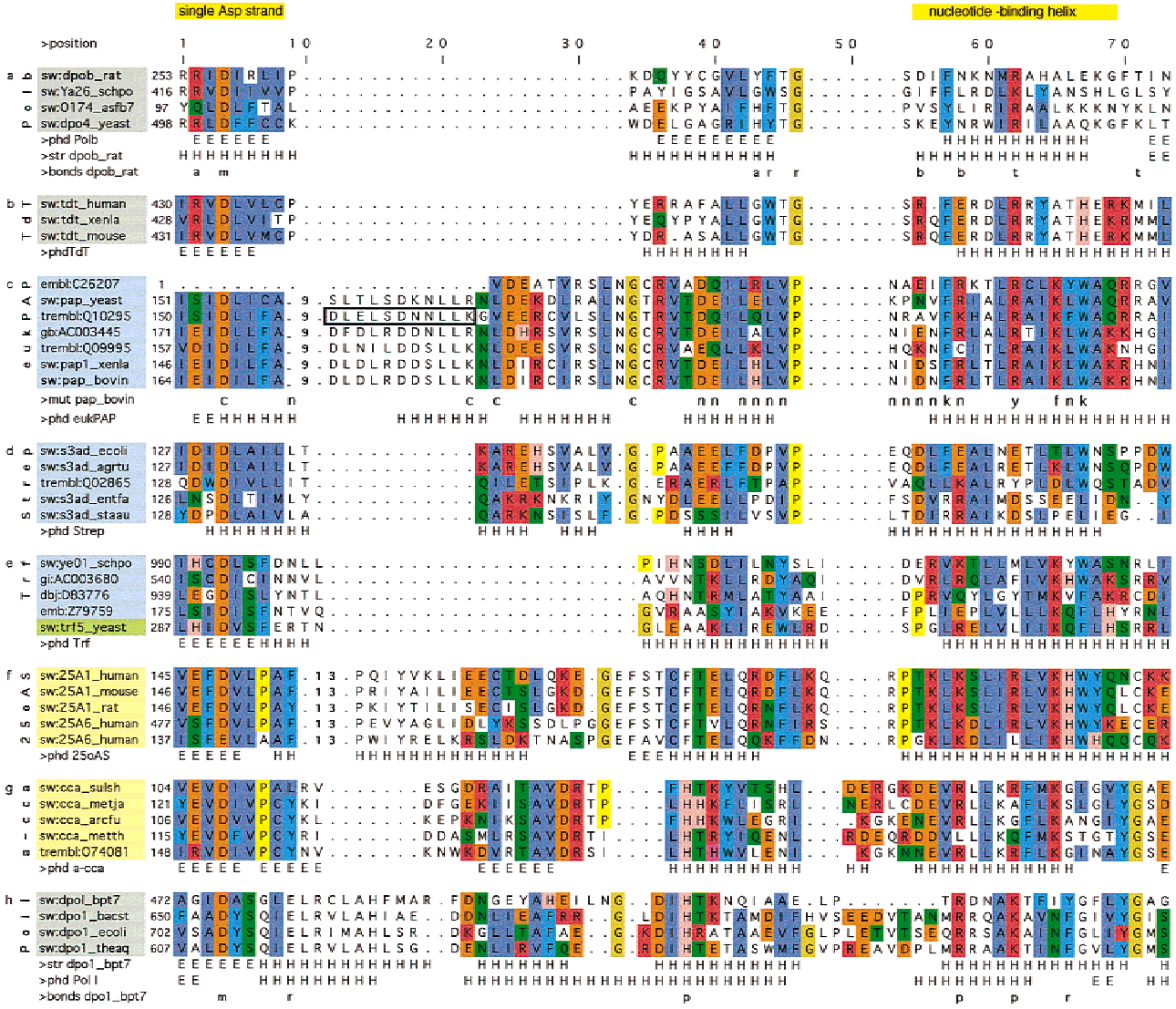

Fig. 5. Multiple sequence alignment of proteins with a homology to PAP. At the top, "single Asp strand" and "nucleotide-binding helix" on yellow background indicate secondary structure elements in Pol β and Pol I that we predict to be present in all the protein families in the alignment. To the left, fields with sequence IDs or accession numbers are colored according to the first (light blue) or second (light yellow) PSI-BLAST search and the overlap (green, see text). Fields with sequence names not detected by the PSI-BLAST are colored light gray. Groups of protein families are: a, DNA polymerase β (pol b, vertical line to the left); b, terminal deoxynucleotidyl transferase (TdT); c, eukaryotic poly(A) polymerases (eukPAP); d, streptomycin adenylyltransferases (Strep); e, topoisomerase-related proteins (Trf); f, 2′-5′ oligo(A) synthetase (2′5′oAS); g, archaeal cca:tRNA nucleotidyltransferases (a-cca); h, Pol I DNA polymerases (PolI). Also on the left, each protein sequence segment is designated by either the Swiss-Prot database ID (sw:), the Trembl database (translated EMBL nucleotide sequence database) accession number (trembl:), the EMBL nucleotide sequence database accession number (embl:) or the genbank database accession number (gb:). The features in the rows below the groups are: >phd, structure prediction (phd-server, EMBL) with the respective sequence group; E in rows signify prediction for extended structure (β-strand) and H for α-helix; >bonds, signifies residues found in hydrogen bonding distance to ribose (r), base (b), template (t), metal (m), phosphate (p), or involved in induced fit mechanism (a); >mut pap_bovin, mutations in bovine PAP (see Table 1) used to measure kinetics parameters and letters in this row indicate: n, not affected; *k*, K_M affected; c, k_{cat} affected; f, both K_M and k_{cat} affected and y, k_{cat} affected if mutated to Tyr. In the sequence rows, amino acids single letter code is on colored background only if conserved 60% or more between neighboring groups, except for Pro and Gly, which are also colored if in adjacent columns to depict parsing. Residues are considered similar if they belong to the group of polar (R, K, D, E, N, Q, S, T, H, Y), hydrophobic (I, F, L, M, C, V, A, W, Y), or aromatic amino acids (F, Y, W). Red background is used for basic residues, orange for acidic residues, green for other hydrophilic residues, cyan for aromatic, pink for His, and blue for hydrophobic amino acids. Yellow background was used with prolines and ochre with glycines. The row of framed fields in the trembl:Q10295 line signifies the tryptic fragment identified to contain a cross-linked N_3ATP in *S. pombe* PAP.

Discussion

The primary aim of this study was to localize the sites in PAP involved in binding the ATP substrate. We used 8-azido-ATP as a probe to identify protein structures in poly(A) polymerases located adjacent to the base of the incoming nucleotide. Addition of Mn^{2+} strongly stimulated cross-linking between N_3ATP and poly(A) polymerases, whereas Mg^{2+} had a relatively small stimulatory effect.

Table 1. *Steady-state kinetics of bovine poly(A) polymerase mutants*

Mutation	re[a]	$K_{M(ATP)}$ (mM)	±[b]	k_{cat} (s^{-1})	±
Wild-type					
bPAP-82kDa	2	0.26	0.03	0.19	0.05
bPAP-513	9	0.24	0.14	0.28	0.14
Helical turn motif					
F100D	2	0.51	0.08	0.0043	0.0011
S102A	2	0.33	0.03	0.37	0.27
Y103G	2	**0.72**	0.02	0.73	0.05
R104A	2	0.45	0.05	**2.66**	0.14
L105A	2	1.04	0.66	1.01	0.10
Prospective fingers domain					
R172G	3	0.30	0.09	0.71	0.32
T177M	2	0.27	0.09	0.46	0.11
K191A	2	0.30	0.01	**0.81**	0.01
L193M	2	0.33	0.02	**1.09**	0.12
G203H	2	0.29	0.04	**0.0007**	0.0004
D208A	2	0.42	0.11	0.31	0.06
E209A	2	0.33	0.03	0.44	0.14
L211K	3	0.33	0.10	0.52	0.12
H212A	2	0.17	0.03	0.53	0.03
L213K	2	0.37	0.04	0.41	0.13
V214Q	3	0.38	0.14	0.27	0.26
N216A	2	0.26	0.03	0.39	0.07
I217T	2	0.25	0.02	0.35	0.05
D218A	2	0.31	0.03	0.41	0.06
N219A	2	0.30	0.04	0.39	0.02
F220R	4	**1.04**	0.30	0.36	0.20
F220S	3	**1.28**	0.71	0.42	0.35
NIDNF216SIANS	2	**3.61**	1.94	0.15	0.05
F220Y	2	0.31	0.05	0.62	0.23
R221A	2	0.37	0.03	0.42	0.23
R225A	5	0.22	0.03	0.19	0.06
R225Y	3	0.42	0.12	**0.84**	0.28
R225N	3	0.15	0.03	0.11	0.02
K228A	3	**1.01**	0.62	**0.09**	0.06
L229R	2	0.46	0.24	0.26	0.18
W230A	3	**1.31**	0.48	0.12	0.07
Y237A	3	**0.88**	0.27	0.16	0.08
F243A	2	**0.77**	0.03	0.21	0.11

[a] re indicates number of determinations.
[b] ± indicates range.

Table 2. *Inhibition of wild-type and mutant PAPs by pyrophosphate*

Mutation	IC_{50}[a] nM/fmol enzyme	±
Wild-type		
bPAP-513	0.82	0.12
Helical turn motif		
F100D	**0.25**	0.03
S102A	**0.32**	0.03
Y103G	0.58	0.03
R104A	0.44	0.01
L105A	0.43	0.04
Prospective fingers domain		
F220R	0.69	0.11
F220S	0.50	0.01
R225A	0.61	0.09
R228A	0.47	0.01
W230A	0.72	0.02
Y237A	0.50	0.02

[a] IC_{50} = concentration at 50% inhibition.

Many examples are known of enzymes requiring divalent metals, including polymerases, which are more active in vitro if Mn^{2+} is substituted for Mg^{2+}. The same is true for poly(A) polymerases, where a lower apparent K_M for the RNA substrate is a reason for the higher activity in the presence of Mn^{2+} (Wahle, 1991). When we measured the K_i for N_3ATP in the presence of Mg^{2+} and Mn^{2+}, we found a 50-fold difference in favor of Mn^{2+}, indicating that N_3ATP binds much better to PAP in the presence of this metal.

Tests for the binding specificity of the N_3ATP photoprobe confirmed a cross-link to a true nucleotide-binding domain. Competition with various ribo- or deoxyribonucleoside triphosphates was in general weaker, but not significantly different from competition with ATP. This indicates that binding of the nucleotides during the early steps of binding occurs predominantly via the phosphate moieties. The conditions for cross-linking were such that a transition state could not be reached because no RNA primer was present. Further tests will be needed to confirm a possible conformational change that may include an induced fit mechanism to increase fidelity in the ATP binding region as the selective event. Such a mechanism has been proposed for Pol β from ternary structure data (Sawaya et al., 1997) and for T7 DNA polymerase, a ternary structure revealed a conformational change in the fingers domain after positioning of the substrates (Doublié et al., 1998).

Cross-linking of catalytic aspartate mutants to N_3ATP confirmed the involvement of these residues in metal binding (Fig. 2C). If Mg^{2+} or Mn^{2+} was included in these reactions, the cross-link was even weaker than with the control. All three aspartate mutants cross-linked very weakly with N_3ATP (approximately 2% compared to wild-type PAP). This is consistent with earlier experiments (Martin & Keller, 1996), which demonstrated that mutation of these aspartate residues to alanine abolished enzyme activity. This indicates that all three aspartates participate in coordination of the two metals. In the Pol β structure (Pelletier et al., 1994), the three catalytic aspartates (residues 190, 192, and 256) were also found to be involved in the coordination of two metal ions. In contrast, two carboxylic acid residues are thought to be sufficient to coordinate the two metals in Pol I polymerases (Brautigam & Steitz, 1998), although mutation of the third carboxylate (E883A) in Klenow polymerase had some effect on catalysis (Polesky et al., 1990). Asp256 in Pol β is needed at one apex of the octahedral coordination shell of metal B, whereas in Pol I-type polymerases the third carboxylic residue was not found to bind the metal ion in site B in the crystal structures (Doublié & Ellenberger, 1998). Our results, therefore, suggest that PAP contains the same metal coordination geometry as Pol β (Fig. 6). Cca:tRNA nucleotidyltransferases could be different in this respect because mutation of an aspartate in the *Sulfolobus shibatae* cca-adding enzyme corresponding to Asp167 in bovine PAP strongly reduced ATP-adding activity but did not affect CTP adding activity (Yue et al., 1998).

Our first attempt to map the location of the cross-linked N_3ATP on the polypeptide was done by CNBr cleavage of the cross-linked

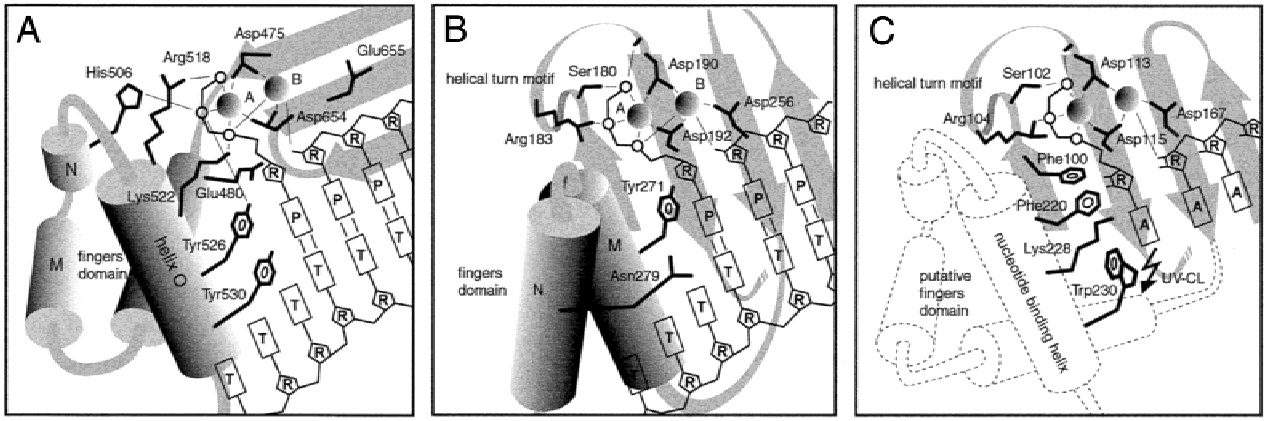

Fig. 6. Comparison of the catalytic core structures of T7 DNA Polymerase, Pol β and a tentative model for the active site of PAP. **A:** T7 DNA polymerase catalytic core. R indicates ribose, B base, and T template. Metals (spheres) are designated A and B. **B:** Pol β catalytic core. **C:** Model for the active site of poly(A) polymerases. The arrow in the lower part indicates the region of the assumed cross-link between azido-ATP and the protein. This model is tentative and is not based on sequence modeling. A denotes adenine bases.

proteins with subsequent analysis by SDS-PAGE. As depicted on the cleavage map (Fig. 3C), the region around amino acid 190 (with respect to bovine PAP) is in almost all cases included in the fragments with the highest amount of label. These results strongly support the idea of a N_3ATP cross-link at about the same site in the bovine PAP as in PAP of *S. pombe*, despite the fact that some variation in the intensity of labeling between newly created subfragments were observed in mutants with methionines introduced in the protein. This is consistent with the idea that this region is near the base of the incoming nucleotide and therefore sensitive to mutagenesis.

Mass spectrometry of tryptic peptides of *S. pombe* PAP allowed the fine mapping of the site of azido-ATP cross-linking. The peptide containing the cross-linked N_3ATP spans from amino acid 167 to amino acid 177. These results, together with those from the CNBr mapping, strongly suggest this region to be near a nucleotide-binding structure.

In the multiple sequence alignment (Fig. 5), stepwise sequence relationships between various types of nucleotidyltransferases are depicted. Iterative PSI-BLAST database searches revealed four protein families with significant homology to the region of the suspected PAP nucleotide binding motif. These enzyme families included streptomycin 3″-adenylyltransferases, the TRF4/5 proteins, the 2′-5′ oligo(A) synthetases, and the archaeal cca:tRNA transferases. The bacterial PAP and ccaT group was found outside the significance threshold. We also included in the alignment the sequence of the fingers domain of several Pol I proteins, which showed a surprising similarity between helix N and O and the region downstream of the single Asp motif in the archaeal ccaTs. The degree of homology and the result of the structure prediction indicates that these two helices are conserved in all nucleotidyltransferase families of the alignment. The motif with the single catalytic aspartate in the Pol I group was aligned with the corresponding motif of Pol β containing the catalytic Asp256. These two motifs have been thought to be homologues (Delarue et al., 1990). However, superimposition of the primer 3′-ends, the metals and the nucleotides of ternary structures of Pol I and Pol β showed that the β-strands of the palm domain in the two enzymes are located almost 90° to one another, whereas the acidic residues superimpose (Steitz et al., 1994). This demonstrated that the two families of enzymes use two different protein architectures to form the same active site geometry for the acidic residues that coordinate the metals. Therefore, according to this structural analogy, Asp705 in *E. coli* Pol I aligns with Asp190 in Pol β (Fig. 6A,B).

The direct molecular link between the β-strand containing the single catalytic Asp in Pol I type polymerases (Asp475 in T7 DNA polymerase) and the fingers domain (helices L, M, N, and O) has the function of a hinge for the movement of the fingers domain during catalysis (Doublié et al., 1998; reviewed in Doublié & Ellenberger, 1998). In addition, the hinge between the β-strand containing catalytic Asp256 and the fingers domain in Pol β was found to take part in an induced fit mechanism (Sawaya et al., 1997). The size of the hinge region in the different protein families is found to be variable according to individual needs.

Although the overall homology between the different groups of proteins in the alignment is moderate, there are some surprising local similarities. Besides the invariant single catalytic Asp at position 4 in Figure 5, there are other conserved features in the region of the nucleotide binding helix. Most striking is the periodic arrangement of the basic Arg and/or Lys residues, which could be expected to be grouped in line on one side of an α-helix. The conserved Phe220 in the eukaryotic poly(A) polymerases could interact with the base or the ribose of the incoming nucleotide. Its mutation to alanine causes an increase of $K_{M(ATP)}$. Strikingly, Phe220 is conserved in TdT, an enzyme involved in elongation of a primer in a template-independent fashion, very much like poly(A) polymerases. It is also possible that residues Phe220, Leu224, and Lys228 of PAP are making contacts with the adenine or the ribose moiety of the incoming nucleotide. These residues would be located on the same face of an α-helix.

The sequence conservation in the alignment between the PAPs and the 2′-5′ oligo(A) synthetase family is also remarkable. Both enzymes polymerize ATP, although they form different types of phosphodiester bonds. Considering the high degree of sequence conservation, residues in the region of the predicted nucleotide binding helix could indeed be responsible for discriminating against bases other than adenine. Moreover, the invariant Trp at position 66 could assist to form a pocket for the nucleotide base in analogy to a tyrosine at the C-terminal end of helix O of many Pol I polymerases (Joyce & Steitz, 1994; Astatke et al., 1998).

As a final test for the postulated nucleotide binding domains in poly(A) polymerases, we analyzed point mutants in the region of the predicted helical turn motif (residues 100 to 105) and fingers domain (amino acids 190 to 250) of bovine PAP and measured their steady-state kinetics parameters. Many mutations had no effect on ATP binding or the catalytic rate of the mutant proteins (Table 1). Hot spots of mutagenic defects were found around amino acid Arg104. Replacement of this large charged side chain resulted in a much higher rate of incorporation of ATP, possibly as a consequence of a missing interaction with ATP that may indirectly constrain catalysis. Another explanation for the high k_{cat} measured with mutant R104A is suggested by the Pol β structure (Sawaya et al., 1997) where the tip of the fingers is in close contact with Arg182 (the analog of Arg104 in PAP). This contact helps to position the fingers domain, which upon closure is moving the helical turn structure toward the phosphoryl oxygens of the dNTPs. If the function of Arg 104 was the same in PAP the mutation R104A could lead to a relaxation of either the selective mechanism for the incoming ATP or it could influence the conformational change within the fingers domain.

Some of the mutants that affected ATP-dependent kinetic parameters in bovine PAP are located in the "nucleotide-binding helix" region (Fig. 5). The effect of the point mutations on the apparent $K_{M(ATP)}$ was moderate (four- to fivefold); only a multiple mutation (NIDNF216 SIANS) increased the $K_{M(ATP)}$ 15-fold. Similar small increases or even decreases in $K_{M(dNTP)}$ were also found in DNA polymerase I when mutants known to interact with nucleotides were tested for their kinetic parameters (Polesky et al., 1990; Astatke et al., 1995; Kaushik et al., 1996). Because several amino acids in the nucleotide binding fold are involved in contacting the incoming nucleotide, the contribution of binding energy for each amino acid is only part of the total contacts.

To identify residues that interact with the oxygens of the triphosphate, we tested inhibition of the activity of PAP mutants with pyrophosphate. Mutants S102A and F100D were more sensitive to pyrophosphate addition when compared to the wild-type. At an analogous position in the Pol β (Ser180) and the KanNt structure (Ser39), serines were found to interact with phosphoryl oxygens of the β and γ phosphates (Pedersen et al., 1995; Sawaya et al., 1997). Thus, the invariant Ser102 in all PAPs is the homolog of the serines found in Pol β and KanNt. Arg104 in PAP is a candidate for interaction with one of the phosphoryl oxygens similar to Arg183 of Pol β, although the positions of these residues are adjacent. Phe100 of PAP is found at an analogous position as Tyr37 in the KanNt structure (Pedersen et al., 1995). There, Tyr37 is located near the ribose of the nucleotide, where it holds the ribose in place. These results further support a role of the helical turn motif in triphosphate binding.

Taken together, our results suggest that poly(A) polymerases contain a catalytic center that strongly resembles that of Pol β (Fig. 6B,C). We assume that the active site also contains a β-sheet structure with three aspartates that coordinate two metal ions. In addition, the helical turn motif as part of the nucleotidyltransferase superfamily signature "G[G/S](x)$_{9-13}$Dx[D/E]" is present in PAPs (Martin & Keller, 1996). Its main function is to bind the triphosphate group of the incoming nucleotide via hydrogen bonds (Fig. 6B,C). Our finding of a N_3ATP cross-link in a region of PAP that is conserved in other nucleotidyltransferase families suggests a role of this region in nucleotide base recognition. It can be anticipated that this region represents the analog to the fingers or nucleotide binding structure found in many polymerases, and that in PAP it has evolved to specifically select for adenine as the base of the incoming nucleoside triphosphate.

Materials and methods

Expression and purification of proteins in E. coli

Expression of poly(A) polymerases was done as described (Martin & Keller, 1996). Proteins expressed in *E. coli* were first affinity purified on Ni-NTA agarose matrix (Qiagen, Hilden, Germany). Fractions containing the protein were further purified on Hitrap heparin and subsequently on Hitrap Q columns (Pharmacia, Uppsala, Sweden).

UV cross-linking of PAP with 8-azido-ATP

For the experiment shown in Figure 1A, the primer $(A)_{15}$ (Xeragon, Zürich, Switzerland) was 5′-end-labeled with [γ-^{32}P]ATP (3000 Ci/mmol, Amersham, Dübendorf, Switzerland) and polynucleotide kinase (Sambrook et al., 1989). Reactions for standard analytical UV cross-linking reactions were done in 0.5 mL thin-wall PCR tubes (Appligene, Gaithersburg, Maryland) and contained in 10 μL: 2.5 μL 4× cross-link buffer (20% glycerol, 40 mM Tris-Cl, pH 8.3, 160 mM KCl), 100–200 ng of PAP, 0.5 mM $MnCl_2$, or 4 mM $MgCl_2$, 0.25 μCi of [γ-^{32}P] 8-azido-ATP (15 to 20 Ci/mmol, ICN Pharmaceuticals, Eschwege, Germany). The reaction was kept on ice for about 2 min, the closed tubes were placed directly on a UV illuminator (Vilber Lourmat, Marne la Vallée, France) with a 302 nm filter and irradiated for 60 s at 180 W. The tubes were put back on ice for 5 min, and the reaction was quenched by the addition of an equal volume of SDS-gel loading buffer (1×: 3% SDS, 1 M β-mercaptoethanol, trace of bromophenolblue). Reactions were heated to 90 °C for 5 min and separated on 10% SDS-polyacrylamide gels. Ribo- and deoxynucleoside triphosphates used for competition experiments were purchased as 100 mM stocks (Boehringer-Mannheim, Mannheim, Germany). Larger scale reactions for CNBr cleavage and preparative tryptic digests were UV cross-linked in aliquots of 50 μL in thin-wall tubes as above. Ten to 200 μg of PAP was mixed at a concentration of 0.5–1.0 μg/μL in cross-link buffer together with 0.5 mM $MnCl_2$, 1 μCi/10 μg of PAP [γ-^{32}P]8-azido-ATP and 0.1 mM of unlabeled 8-azido-ATP (Sigma, St. Louis, Missouri) was included in the reaction.

Preparative reversed-phase chromatography and mass spectral analysis

After UV cross-linking, proteins were reduced and carboxymethylated. For this, 0.6 w/v of urea was added to the reaction. After addition of Tris-Cl, pH 8.0 to 0.1 M, EDTA to 10 mM, and DTT to 10 mM, the reaction was heated to 37 °C for 1 h. Iodacetamide (Fluka, Buchs, Switzerland) was added to 20 mM, and the reaction was incubated further for 15 min in the dark. For desalting and removal of free label, the protein was loaded on a 10 mm × 20 cm Sephadex G-50 column, equilibrated in 0.1 mM NH_4-bicarbonate, and eluted with the same buffer. Poly(A) polymerase was found to be resistant to cleavage by trypsin. The protein was therefore precleaved with endoproteinase LysC (*Achromobacter* protease I, from WACO) in the presence of 6 M urea. For this, fractions containing the labeled protein from the Sephadex column were pooled and

dried down to about 25 μL. Three volumes of 8 M urea in 0.1 M NH_4-bicarbonate were added, and the mixture was heated to 37 °C for 10 min. After addition of 2 μg of protease Lys C, the reactions were incubated for 2 h at 37 °C. The resulting digest was subsequently recleaved with trypsin (Promega, Madison, Wisconsin) after lowering the urea concentration to 2 M or below. For this, 200 μL of NH_4-bicarbonate and 2–5 μg of trypsin was added, and the digestion was continued overnight at 37 °C. The digested material was directly loaded on a 1 mL RESOURCE-RPC column (Pharmacia) for pre-fractionation and eluted with a gradient from 0 to 80% solvent B [80% $CNCH_3$ in 0.1% trifluoroacetic acid (TFA)] vs. solvent A (0.1% TFA). One milliliter fractions were collected and fractions containing labeled material were determined by Cerenkov counting. Fractions of interest were pooled, dried in a Speed-Vac evaporator, resuspended in 5% CH_3CN in 0.1% TFA, and loaded on a 2.1 mm VYDAC C_{18} column (Vydac, The Separations Group, Hesperia, California). The same gradient of solvent B was applied as in the previous column and fractions of 50 μL were collected. Fractions containing labeled material were dried and resuspended in 0.1% acetic acid and 50% methanol. Aliquots were injected into a TSQ 7000 mass spectrometer (Finnigan, San José, California), equipped with a micro electrospray ionization source (Davis et al., 1995).

CNBr digestion of PAP, gel electrophoresis, and blotting to PVDF membranes

For CNBr cleavage, 10 or 20 μg of PAP, cross-linked with N_3ATP, was precipitated with four volumes of cold acetone and resuspended in 70% formic acid. An equal volume of 200 mg/mL of CNBr (Kodak) in 70% formic acid (fresh or as stock kept at −80 °C) was added, and the reaction was incubated overnight at room temperature in the dark. The reactions were dried in a Speed-Vac, resuspended in 300 μL of H_2O and redried. The pellet was resuspended in Tricine-gel loading buffer, heated to 37 °C and loaded on a 16.5% Tricine gel (Schägger & von Jagow, 1987). Gels were blotted to PVDF membranes (Immobilon-P, Millipore, Bedford, Massachusetts) and exposed to a PhosphorImager screen.

Mutagenesis, steady-state kinetics, and pyrophosphate inhibition measurements

Site-directed point mutations D113A, D115A, D167A, F100R, and Y103A in the bPAP-513 recombinant clone were described (Martin & Keller, 1996). Other mutants were generated as described (Mikaelian & Sergeant, 1992). Kinetic parameters K_M and k_{cat} for ATP were determined in the presence of Mg^{2+}. The reaction was done in 25 μL volume, containing 10% glycerol, 25 mM Tris, pH 8.3, 40 mM KCl, 4 mM $MgCl_2$, 0.1 mg/mL BSA (Boehringer-Mannheim), 0.01% NP40, 7 μg of poly(A) (Boehringer-Mannheim), 40 to 100 ng of PAP and 0.25 μCi of [α-^{33}P]ATP (2,000 Ci/mmol, from NEN, Boston, Massachusetts). ATP was titrated between 0.04 and 0.6 mM. Reactions were incubated at 37 °C for 20 min and stopped by adsorption to DE-81 filters as described (Martin & Keller, 1996). Pyrophosphate inhibition was tested by titration of pyrophosphate between 0.1 and 0.8 mM into reactions containing 10% glycerol, 25 mM Tris-Cl, pH 8.3, 40 mM KCl, 0.1 mg/mL BSA, 0.01% NP-40, 0.5 mM $MnCl_2$, 0.5 mM ATP, 0.25 μCi [α-^{33}P]ATP, 1 μg $(A)_{15}$, and 10 ng of PAP and the reactions were incubated for 15 min at 37 °C. Reactions were absorbed to DE-81 filters as in the kinetics assays. IC_{50} (concentration of PP_i at 50% inhibition) was calculated with Microsoft excel.

Acknowledgments

We thank Thierry Mini for excellent technical assistance in mass spectrometry and Diana Blank for help with the artwork. We thank Lionel Minvielle-Sebastia and Martin Ohnacker for reading the manuscript, and we are indebted to Sylvie Doublié and Elmar Wahle for helpful discussions. This work was supported by the Kantons of Basel, the Swiss National Science Foundation and the European Union via the Bundesamt für Bildung und Wissenschaft (BBW), Bern, and the Louis Jeantet Foundation for Medicine.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res 25*:3389–402.

Aravind L, Koonin EV. 1999. DNA polymerase β-like nucleotidyltransferase superfamily: Identification of three new families, classification and evolutionary history. *Nucleic Acids Res 27*:1609–1618.

Astatke M, Grindley NDF, Joyce CM. 1995. Deoxynucleotide triphosphate and pyrophosphate binding sites in the catalytically competent ternary complex for the polymerase reaction catalyzed by DNA polymerase I (Klenow fragment). *J Biol Chem 270*:1945–1954.

Astatke M, Grindley ND, Joyce CM. 1998. How *E. coli* DNA polymerase I (Klenow fragment) distinguishes between deoxy- and dideoxynucleotides. *J Mol Biol 278*:147–165.

Ballantyne S, Bilger A, Åström J, Virtanen A, Wickens M. 1995. Poly(A) polymerases in the nucleus and cytoplasm of frog oocytes: dynamic changes during oocyte maturation and early development. *RNA 1*:64–78.

Brautigam CA, Steitz TA. 1998. Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr Opin Struct Biol 8*:54–63.

Cao G, Sarkar N. 1992. Identification of the gene for an *Escherichia coli* poly(A) polymerase. *Proc Natl Acad Sci USA 89*:10380–10384.

Castano IB, Brzoska PM, Sadoff BU, Chen H, Christman MF. 1996. Mitotic chromosome condensation in the rDNA requires TRF4 and DNA topoisomerase I in *Saccharomyces cerevisiae*. *Genes Dev 10*:2564–2576.

Cheng N, Merrill BM, Painter GR, Frick LW, Furman PA. 1993. Identification of the nucleotide binding site of HIV-1 reverse transcriptase using dTTP as a photoaffinity label. *Biochemistry 32*:7630–7634.

Colgan DF, Manley JL. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes Dev 11*:2755–2766.

Davies JF, Almassy RJ, Hostomska Z, Ferre RA, Hostomsky Z. 1994. 2.3 Å crystal structure of the catalytic domain of DNA polymerase β. *Cell 76*:1123–1133.

Davis MT, Stahl DC, Hefta SA, Lee TD. 1995. A microscale electrospray interface for on-line, capillary liquid chromatography/tandem mass spectrometry of complex peptide mixtures. *Anal Chem 67*:4549–4556.

Delarue M, Poch O, Tordo N, Moras D, Argos P. 1990. An attempt to unify the structure of polymerases. *Protein Eng 3*:461–467.

Doublié S, Ellenberger T. 1998. The mechanism of action of T7 DNA polymerase. *Curr Opin Struct Biol 8*:704–712.

Doublié S, Tabor S, Long AM, Richardson CC, Ellenberger T. 1998. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature 391*:251–258.

Edmonds M. 1982. Poly(A) adding enzymes. In: Boyer PD, ed. *The enzymes, Vol XV*. New York: Academic Press, Inc. pp 217–240.

Evans RK, Beach CM, Coleman MS. 1989. Photoaffinity labeling of terminal deoxynucleotidyl transferase. 2. Identification of peptides in the nucleotide binding domain. *Biochemistry 28*:713–720.

Gardner A. 1995. TREMBL database accession number: Q09995. Hinxton, U.K.: European Bioinformatics Institute, EMBL outstation.

Gebauer F, Richter JD. 1995. Cloning and characterization of a *Xenopus* poly(A) polymerase. *Mol Cell Biol 15*:1422–1430.

Gershon PD, Ahn B, Garfield M, Moss B. 1991. Poly(A) polymerase and dissociable polyadenylation stimulatory factor encoded by vaccinia virus. *Cell 66*:1269–1278.

Hadidi A, Sethi S. 1976. Polyadenylate polymerase from cytoplasm and nuclei of N.I.H.-Swiss mouse embryos. *Biochim Biophys Acta 425*:95–109.

Holm L, Sander C. 1995. DNA polymerase β belongs to an ancient nucleotidyltransferase superfamily. *Trends Biochem Sci 20*:345–347.

Ingle CA, Kushner SR. 1996. Development of an *in vitro* mRNA decay system for *Escherichia coli*: Poly(A) polymerase is necessary to trigger degradation. *Proc Natl Acad Sci USA 93*:12926–12931.

Ishii N, Aoki Y, Arisawa M. 1997. GenBank database accession number AB009394. Bethesda, Maryland: National Center for Biotechnology Information, NIH.

Joyce CM, Steitz TA. 1994. Function and structure relationships in DNA polymerases. *Annu Rev Biochem 63*:777–822.

Kaushik N, Pandey VN, Modak MJ. 1996. Significance of the O-helix residues of *Escherichia coli* DNA polymerase I in DNA synthesis: Dynamics of the dNTP binding pocket. *Biochemistry 35*:7256–7266.

Keller W, Minvielle-Sebastia L. 1997. A comparison of mammalian and yeast pre-mRNA 3′-end processing. *Curr Opin Cell Biol 9*:329–336.

Kiefer JR, Mao C, Hansen CJ, Basehore SL, Hogrefe HH, Braman JC, Beese LS. 1997. Crystal structure of a thermostable *Bacillus* DNA polymerase I large fragment at 2.1 Å resolution. *Structure 5*:95–108.

Li Y, Korolev S, Waksman G. 1998. Crystal structure of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: Structural basis for nucleotide incorporation. *EMBO J 17*:7514–7525.

Lingner J, Kellermann J, Keller W. 1991. Cloning and expression of the essential gene for poly(A) polymerase from *S. cerevisiae*. *Nature 354*:496–498.

Manley JL, Takagaki Y. 1996. The end of the message-another link between yeast and mammals. *Science 274*:1481–1482.

Martin G, Keller W. 1996. Mutational analysis of mammalian poly(A) polymerase identifies a region for primer binding and a catalytic domain, homologous to the family X polymerases, and to other nucleotidyltransferases. *EMBO J 15*:2593–2603.

Mikaelian I, Sergeant A. 1992. A general and fast method to generate multiple site directed mutations. *Nucleic Acids Res 20*:376.

Nikodem V, Fresco JR. 1997. Protein fingerprinting by SDS-gel electrophoresis after partial fragmentation with CNBr. *Anal Biochem 97*:382–386.

Ohnacker M, Minvielle-Sebastia L, Keller W. 1996. The *Schizosaccharomyces pombe pla1* gene encodes a poly(A) polymerase and can functionally replace its *Saccharomyces cerevisiae* homologue. *Nucleic Acids Res 24*:2585–2591.

Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA. 1985. Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature 313*:762–766.

Pedersen LC, Benning MM, Holden HM. 1995. Structural investigation of the antibiotic and ATP-binding sites in kanamycin nucleotidyltransferase. *Biochemistry 34*:13305–13311.

Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J. 1994. Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer and ddCTP. *Science 264*:1891–1903.

Plough M, Jensen AL, Barkholt V. 1989. Determination of amino acid compositions and NH_2-terminal sequences of peptides electroblotted onto PVDF membranes from tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis: Application to peptide mapping of human complement C3. *Anal Biochem 181*:33–39.

Polesky AH, Steitz TA, Grindley NDF, Joyce CM. 1990. Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from *Escherichia coli*. *J Biol Chem 265*:14579–14591.

Potter RL, Haley BE. 1983. Photoaffinity labeling of nucleotide binding sites with 8-azidopurine analogs: Techniques and applications. *Methods Enzymol 91*:613–633.

Raabe T, Bollum FJ, Manley JL. 1991. Primary structure and expression of bovine poly(A) polymerase. *Nature 353*:229–234.

Rost B, Sander C. 1993. Prediction of protein structure at better than 70% accuracy. *J Mol Biol 232*:584–599.

Rush J, Konigsberg WH. 1990. Photoaffinity labelling of the Klenow fragment with 8-azido-dATP. *J Biol Chem 265*:4821–4827.

Sakon J, Liao HH, Kanikula AM, Benning MM, Rayment I, Holden HM. 1993. Molecular structure of kanamycin nucleotidyltransferase determined to 3.0-Å resolution. *Biochemistry 32*:11977–11984.

Salvucci ME, Chavan AJ, Haley BE. 1992. Identification of peptides from the adenine binding domains of ATP and AMP in adenylate kinase: Isolation of photoaffinity-labeled peptides by metal chelate chromatography. *Biochemistry 31*:4479–4487.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning. A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sawaya MR, Pelletier H, Kumar A, Wilson SH, Kraut J. 1994. Crystal structure of rat DNA polymerase β: Evidence for a common polymerase mechanism. *Science 264*:1930–1935.

Sawaya MR, Prasad R, Wilson SH, Kraut J, Pelletier H. 1997. Crystal structures of human DNA polymerase β complexed with gapped and nicked DNA: Evidence for an induced fit mechanism. *Biochemistry 36*:11205–11215.

Schägger H, von Jagow G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem 166*:368–379.

Srivastava DK, Evans RK, Kumar A, Beard WA, Wilson SH. 1996. dNTP binding site in rat DNA polymerase β revealed by controlled proteolysis and azido photoprobe cross-linking. *Biochemistry 35*:3728–34.

Steitz TA, Smerdon SJ, Jäger J, Joyce CM. 1994. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science 266*:2022–2025.

Tabor S, Richardson CC. 1995. A single residue in DNA polymerases of the *Escherichia coli* DNA polymerase I family is critical for distinguishing between deoxy- and dideoxyribonucleotides. *Proc Natl Acad Sci USA 92*:6339–6343.

Thuresson A, Åström J, Åström A, Grönvik K, Virtanen A. 1994. Multiple forms of poly(A) polymerases in human cells. *Proc Natl Acad Sci USA 91*:979–983.

Wahle E. 1991. Purification and characterization of a mammalian polyadenylate polymerase involved in the 3′ end processing of messenger RNA precursors. *J Biol Chem 266*:3131–3139.

Wahle E, Keller W. 1996. The biochemistry of polyadenylation. *Trends Biochem Sci 21*:247–250.

Wahle E, Martin G, Schiltz E, Keller W. 1991. Isolation and expression of cDNA clones encoding mammalian poly(A) polymerase. *EMBO J 10*:4251–4257.

Wahle E, Rüegsegger U. 1999. 3′-end processing of pre-mRNA in eukaryotes. *FEMS Microbiol Rev 23*:277–295.

Yamamoto K, Sasaki T. 1997. GenBank database accession number C26207. Bethesda, Maryland: National Center for Biotechnology Information, NIH.

Yue D, Maizels N, Weiner AM. 1996. CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: Characterization of the CCA-adding enzyme from the archæal hyperthermophile *Sulfolobus shibatae*. *RNA 2*:895–908.

Yue D, Weiner AM, Maizels N. 1998. The CCA-adding enzyme has a single active site. *J Biol Chem 273*:29693–700.