MOLECULAR AND CELLULAR BIOLOGY, May 1994, p. 2946–2957
0270-7306/94/$04.00+0

Vol. 14, No. 5

Poly(A) Polymerase Contains Multiple Functional Domains

TOBIAS RAABE,† K. G. K. MURTHY, AND JAMES L. MANLEY*

Department of Biological Sciences, Columbia University, New York, New York 10027

Received 14 September 1993/Returned for modification 17 November 1993/Accepted 24 January 1994

Poly(A) polymerase (PAP) contains regions of similarity with several known protein domains. Through site-directed mutagenesis, we provide evidence that PAP contains a functional ribonucleoprotein-type RNA binding domain (RBD) that is responsible for primer binding, making it the only known polymerase to contain such a domain. The RBD is adjacent to, and probably overlaps with, an apparent catalytic region responsible for polymerization. Despite the presence of sequence similarities, this catalytic domain appears to be distinct from the conserved polymerase module found in a large number of RNA-dependent polymerases. PAP contains two nuclear localization signals (NLSs) in its C terminus, each by itself similar to the consensus bipartite NLS found in many nuclear proteins. Mutagenesis experiments indicate that both signals, which are separated by nearly 140 residues, play important roles in directing PAP exclusively to the nucleus. Surprisingly, basic amino acids in the N-terminal-most NLS are also essential for AAUAAA-dependent polyadenylation but not for nonspecific poly(A) synthesis, suggesting that this region of PAP is involved in interactions both with nuclear targeting proteins and with nuclear polyadenylation factors. The serine/threonine-rich C terminus is multiply phosphorylated, including at sites affected by mutations in either NLS.

Poly(A) polymerase (PAP) plays an important role in the 3′-end processing of eukaryotic pre-mRNAs (for reviews, see references 38, 66, and 68). 3′-end processing is best understood in mammals and occurs in two steps that can be accurately reproduced in vitro. The first step is cleavage of the pre-mRNA at a defined site, usually 10 to 30 nucleotides downstream of the AAUAAA polyadenylation signal, and the second is synthesis of the poly(A) tail on the newly generated RNA 3′ terminus. The two steps are normally tightly coupled but can be uncoupled experimentally in vitro (37, 42, 71). While multiple protein factors are required for the complete reaction (15, 16, 23, 59, 60), only two are essential for the poly(A) addition phase, PAP and the multisubunit cleavage-polyadenylation specificity factor (CPSF) (6, 10, 44, 60, 62, 65). PAP cannot by itself recognize the AAUAAA signal. Instead, CPSF, which consists of 160-, 105-, 70-, and possibly 35-kDa polypeptides, does so by directly and specifically binding to the AAUAAA (7, 10, 27). RNA gel shift analyses indicate that PAP and CPSF can interact cooperatively to form an RNA-protein complex with AAUAAA-containing RNA (44). In addition, PAP can nonspecifically polyadenylate essentially any RNA 3′ terminus regardless of its sequence; this nonspecific reaction is greatly stimulated by manganese (reviewed in reference 20).

The recent cloning and expression of cDNAs encoding bovine (48, 67) and yeast (35) PAPs have made it possible to begin to understand the molecular basis of the nonspecific and specific activities of PAP. Two related PAP cDNAs were isolated from a bovine cDNA library (48). They encode the nearly identical proteins PAP I and PAP II, of 77 and 83 kDa, respectively, which differ only at their C termini and are presumably generated by alternative splicing. cDNAs capable of encoding a 43-kDa protein (PAP III), which appears to be a truncated form of PAP that also arises from alternative splicing, have also been isolated (48a, 67). The N-terminal halves of yeast and bovine PAPs are 47% identical (35), indicating that PAP is highly conserved throughout evolution.

In bovine PAP, several similarities to consensus sequences of known protein domains have been noted (48). These include a ribonucleoprotein (RNP)-type RNA binding domain (RBD [5]), or RNA recognition motif (47), which was first detected in the yeast poly(A)-binding protein (1, 54). Another is the so-called polymerase module (PM [3]), which represents a part of the catalytic domain of all known RNA-dependent polymerases (46). The putative RBD and PM domains overlap and are located in the N-terminal half of the protein. Toward the C terminus, a possible bipartite nuclear localization signal (NLS [reviewed in reference 18]) is immediately followed by a long serine/threonine (S/T)-rich region. This S/T-rich region is not essential for either specific or nonspecific polyadenylation in vitro but contains many consensus phosphorylation sites and thus may have a possible regulatory function (48).

Similarities to the RBD consensus have been observed in more than 40 known or suspected RNA-binding proteins (19, 28). With the possible exception of *Escherichia coli* rho factor (see reference 28 for a discussion), no enzymes that contain an RBD motif have been described; RBD-containing proteins all appear to be RNA-binding proteins without obvious enzymatic activities. Different RBDs have strikingly different preferences for specific RNA sequences or structures [e.g., the poly(A)-binding protein for poly(A) and the U1 small nuclear RNP 70K protein for a specific stem-loop structure]. The RBD consists of about 80 amino acids and contains two characteristic elements, a highly conserved RNP consensus sequence of eight residues (RNP1) and another conserved sequence of six residues (RNP2). The structure of the RBD contains a solvent-exposed four-stranded β sheet supported by two α helices, as determined by crystallographic (45) and nuclear magnetic resonance (24) studies of the U1A small nuclear RNP protein and the heterogeneous nuclear RNP C protein (69). In this structure, RNP1 and RNP2 are juxtaposed on adjacent β strands, and site-directed mutagenesis has shown that they are important for RNA binding (45). Two highly conserved aromatic residues at positions 3 and 5 and one basic residue at position 1 in the RNP1 octamer and an aromatic residue in

* Corresponding author.
† Present address: The Picower Institute for Medical Research, Manhasset, NY 11030.

TABLE 1. Sequences of oligonucleotides used to generate mutant PAPs[a]

Mutant PAP	Sequence
RNP2 (LI → SR)	5′-TAATTTTCCCAATCTAGAAATCCTGCG-3′
RNP1	
K-1 → H	AATGTAAAAATATGGCCACCAACAT
K-1 → N	AATGTAAAAATATTGCCACCAACAT
F-3 → N	AGATCCAAAAGTATTAATTTTCCCA
F-3 → S	AGATCCAAAAGTACTAATTTTCCCA
F-5 → I	TATAAGATCCGATCGTAAAAATTTT
F-5 → A	TATAAGATCCGGCCGTAAAAATTTT
FTF → SRS	ATAAGATCCAGATCTAGAAATTTTCCC
Motif A	
ΔVD-D	GGTGAAAAAAGATCTATGTCTTGGT
D5 → S	AATCACTTCGCGAAACATGTCTT
D5 → A	AAAATCACTTCGCGCAACATGTCT
D8 → A	GGTGAAAAAAGCGCTCCGATCAACA
SD → AA	GAGGTGAAAAACGCGGCTCGATCAACA
Motif B (DG → AA)	CAATCTCTATCGCGGCAAAACACAG
Motif C	
DD → NN	TAAGCAGACTGTTATTACGTAAGTCCAAG
DD → AA	TAAGCAGACTCGCGGCTCTTAAGAC
NLS1A (KKKK → NNIK)	TTCTGTTGAATGAATATTATTTTTCTGAAG
NLS1B (KRK → NLN)	ATGGAGTTGGTTTAAATTTACATGCAT
NLS2 (KKTK → NITK)	CTGTTTTGGTAATATTGGGGCTC

[a] The indicated motifs are shown in Fig. 1 and described in the text.

position 2 of the RNP2 hexamer may directly interact with single-stranded RNA (40, 45; reviewed in reference 28).

Similarity to the PM in PAP is found beginning at the C terminus of the putative RBD (48). The PM consists of four poorly conserved consensus elements, termed motifs A to D, that are spaced rather variably but apparently arranged in a common tertiary structure (30). Substitutions of 100% conserved aspartic acid residues in motif A or C have been shown to block the polymerization activities in three RNA-dependent polymerases, human immunodeficiency virus (HIV) reverse transcriptase (12, 33, 36) and the L-A virus (50) and Bunyamwera virus (26) RNA polymerases. According to the HIV reverse transcriptase crystal structure, these residues may bind catalytic divalent metal ions (30).

PAP has long been suspected to be a nuclear protein, as its first suggested in vivo function, polyadenylation of pre-mRNAs, is a nuclear process, and the presence of sequence similarity in PAP to the bipartite NLS (52) is consistent with this. However, polyadenylation of maternally stored mRNAs has been shown to occur in the cytoplasm of enucleated *Xenopus* oocytes (reviewed in references 25, 51, and 68), and a large fraction of PAP activity can be detected in HeLa cytoplasmic extracts (53a). It is currently unclear whether cytoplasmic polyadenylation is a universal process or occurs only early in development and whether a different or the same PAP is responsible for nuclear and cytoplasmic polyadenylation.

The sequence motifs observed in PAP are intriguing, but their functional significance is unknown. On the basis of the effects of amino acid substitutions and deletions, we provide evidence that PAP indeed contains an RNP-type RBD that appears to partially overlap a catalytic domain that is related to, but distinct from, a classical PM. We also show that PAP I and PAP II are exclusively nuclear, while the truncated form, PAP III, is predominantly cytoplasmic. However, PAP I becomes cytoplasmic if, in addition to the bipartite NLS, a second NLS within the S/T-rich region is also mutated. While the first NLS is thus not the sole mediator of nuclear localization, mutations in this motif were unexpectedly found to abolish AAUAAA-dependent polyadenylation in vitro but to have no effect on nonspecific poly(A) synthesis.

MATERIALS AND METHODS

Construction of mutant PAP cDNAs. The oligonucleotides shown in Table 1 were used for site-directed mutagenesis. Single-stranded DNA from phagemid pBS T1NE PAP I (termed PAP 4 in reference 48) was isolated from *E. coli* CJ 236, and mutagenesis was carried out by the Bio-Rad protocol. The resulting mutant plasmids were used to generate templates for in vitro transcription as described previously (48). Internal deletion mutations were constructed by standard procedures using convenient restriction sites. C-terminal truncations were generated by treatment with exonuclease III and nuclease S1 as described by the supplier (Promega). Nucleotide sequences were verified by DNA sequencing, and the resulting plasmids were used as templates for PCR. Mammalian expression vectors containing PAP cDNAs fused at the PAP start codon to the influenza virus hemagglutinin (flu) epitope (21) were constructed by cloning several of the pBS T1NE PAP constructs described above into the pCGN vector (61).

In vitro translation of wild-type and mutant PAPs. RNAs transcribed from wild-type and mutant templates were phenol extracted, ethanol precipitated, and used in in vitro translation reactions at a concentration of 20 μg/ml, using RNase-treated reticulocyte lysate (Promega) as suggested by the supplier. To monitor the efficiency of translation, [^{3}H]leucine (Amersham) was added to reaction mixtures, and proteins were resolved by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and visualized by fluorography.

Polyadenylation assays. Specific and nonspecific polyadenylation reactions were carried out essentially as described by Raabe et al. (48). For both assays, 2 μl of in vitro-translated PAP or 1 μl of HeLa PAP (53a) was used. For the specific assay, PAP was incubated together with 4 μl of bovine CPSF (44) and an SP6 polymerase-produced 180-nucleotide-long simian virus 40 late pre-mRNA that terminated 6 nucleotides after the AAUAAA (53). Incubation was for 1.5 h at 30°C as described previously (48). Conditions for the nonspecific assay, which used an oligo(A_5) primer and was done in the presence of 0.5 mM $MnCl_2$, were exactly as described previously (48) except that in the experiment shown in Fig. 5, the primer concentration was varied as indicated. All experiments were repeated a minimum of three independent times. For the specific assay, typical results are shown, while results from the nonspecific assays were averaged.

Transient transfection and immunostaining of wild-type and mutant PAPs. For transfections of 293 human kidney cells for immunofluorescence, 2 μg of the appropriate flu epitope-tagged PAP cDNA constructs was added to 60-mm-diameter plates of about 30% confluent cells, using the calcium phosphate method. A glycerol shock was applied after overnight incubation, and the cells were further grown for a total of 48 h after addition of DNA. For indirect immunofluorescence microscopy, cells were washed with phosphate-buffered saline (PBS) twice for 5 min each time and fixed with 3.7% formaldehyde for 20 min at room temperature. After two 5-min washes with PBS, the cells were permeabilized with 1% (vol/vol) Nonidet P-40 in PBS for 20 min. After four additional short PBS washes, 50 μl of 1:10-diluted monoclonal antibody 12CA5 (21) in 2% bovine serum albumin (BSA)–10% calf

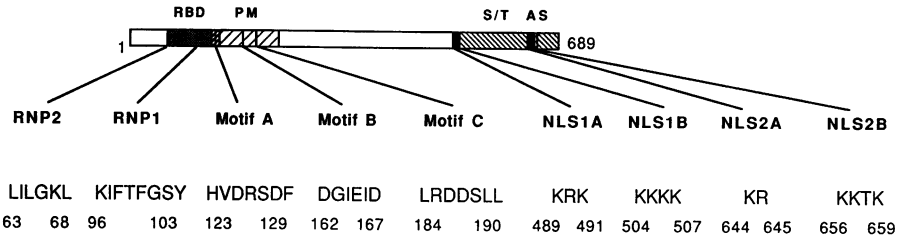

FIG. 1. Possible functional domains of PAP. The putative domains of PAP, with their most conserved sequence elements, are shown. The RBD contains the two highly conserved elements, RNP1 and RNP2. The PM contains four motifs, A to D, one of which (D) is not shown. The upstream bipartite NLS (NLS1) contains two clusters of basic residues, here termed NLS1A and NLS1B. The C-terminal-most bipartite NLS (NLS2) also contains two basic clusters, here termed NLS2A and NLS2B, and is located within the S/T-rich C terminus just N terminal to the point of sequence divergence arising from alternative splicing (AS). Conserved amino acid residues are indicated by the single-letter code, and the positions in the PAP sequence of the first and last residues in each motif are given as specified in reference 48.

serum–PBS was added; the mixture was incubated for 45 min at 37°C under coverslips and then washed four times with PBS. The mouse monoclonal antibody was probed with fluorescein-labeled anti-mouse immunoglobulin G secondary antibody (1:50 dilution in 2% BSA–10% calf serum–PBS; Cappel) at 37°C for 45 min as described above. Cells were washed four times with PBS and mounted.

Western blotting (immunoblotting). For Western blotting, 100-mm-diameter plates of 293 cells were each transfected with 6 μg of the appropriate PAP expression construct. Cells were washed in PBS, and cells from two plates were resuspended in 300 μl of extraction buffer (50 mM Tris [pH 6.8], 1 mM dithiothreitol, 0.2 mM phenylmethylsulfonyl fluoride). Cells were then lysed by sonication, diluted with 2× sample buffer, boiled, and fractionated on an SDS–10% polyacrylamide gel (32). Proteins were then electroblotted (63) onto nitrocellulose BA85 (Schleicher & Schuell) at 200 mA for 2 h. The nitrocellulose membrane was probed with monoclonal antibody 12CA5 (21), using alkaline phosphatase-conjugated anti-mouse immunoglobulin G (Sigma) as a secondary antibody. The presence of two cross-reacting endogenous proteins (see Fig. 9) was used to verify that equal cell equivalents were analyzed in each sample.

Phosphatase treatment of PAP. 293 cells were transfected with the flu epitope-tagged PAP I expression construct and processed as for Western blotting except that the extraction buffer contained 100 mM NaCl, 50 mM piperazine-*N*,*N*′-bis (2-ethanesulfonic acid) (PIPES; pH 6.0), 1 mM dithiothreitol, and 0.2 mM phenylmethylsulfonyl fluoride. Ten-microliter portions of the supernatant obtained following a brief centrifugation in a microcentrifuge were incubated in the presence of the phosphatase inhibitor NaF (50 mM, final concentration) plus the amount of potato acid phosphatase (Boehringer) indicated in Fig. 9, after addition of water to a 40-μl final volume. Incubation was for 30 min at 37°C. Reaction products were separated by SDS-PAGE and analyzed by Western blotting using the anti-flu epitope monoclonal antibody 12CA5.

RESULTS

Conserved residues in the putative RBD are essential for polyadenylation. The region of PAP displaying similarity to the RBD consensus lies between residues 62 and 139 (Fig. 1). The RNP1 octamer is the more conserved of the two RBD elements, and within RNP1, a basic residue in position 1 (in PAP a lysine) and aromatic residues in positions 3 and 5 (in PAP phenylalanines) are very highly conserved. The basic residue is required for RNA binding in the U1A protein (45), and substitution of phenylalanine at position 3 or 5 significantly reduced RNA binding by *E. coli* rho factor (13). On the other hand, a similar mutation of a phenylalanine at position 5 in a yeast poly(A)-binding protein RBD was without detectable effect (55).

To test the functional significance of these similarities in PAP, a number of mutations were introduced into the RNP1 and RNP2 motifs (see Fig. 3A) as described in Materials and Methods. The resulting mutant PAP DNAs were used as templates for in vitro transcription, and the RNAs produced were translated in reticulocyte lysate. Expression levels of all mutant PAPs were similar or identical to that of wild-type PAP (Fig. 2). These PAP preparations were then used to measure activity, in both specific AAUAAA-dependent and nonspecific poly(A) synthesis assays, as described previously (48). Specific polyadenylation was measured by incubating wild-type or mutant PAPs with highly purified bovine CPSF (44) and a ^{32}P-labeled simian virus 40 late pre-mRNA (Fig. 3C). Nonspecific poly(A) synthesis activity was measured in the presence of $MnCl_2$ by incorporation of [α-^{32}P]ATP onto an AAAAA primer (Fig. 3B). While in certain instances the results described below revealed differences in the requirement of specific PAP residues in the two assays, in general the mutants behaved similarly in both. In these cases, the nonspecific assay was more sensitive to small differences and could be more readily quantitated. Indeed, we did not attempt to quantitate the results of the specific assay rigorously and relied on visual inspection of multiple autoradiograms. Conditions for both assays were adjusted so that the concentration of PAP was rate limiting, and assays were repeated multiple times with at least three completely independent preparations of the various PAP derivatives.

Because the findings mentioned above raised the possibility that single amino acid substitutions have limited effects, we first constructed and analyzed a triple substitution changing the phenylalanines in positions 3 and 5 of RNP1 to serines and the threonine in position 4 to arginine (FTF→SRS). As shown in Fig. 3, this mutation almost completely inhibited both nonspecific and specific polyadenylation. (The slight extension of the primer in the specific assay resulted from endogenous PAP activity in the reticulocyte lysate; compare the last two lanes in Fig. 3C; see also reference 48.) Single mutations of each phenylalanine were then constructed and analyzed. F-3→S or F-3→N and F-5→A mutations all reduced, but did not eliminate, both nonspecific and specific activity (although the effect of F-3→S in the specific assay was slight). The substitution F-5→I was unusual in that it increased the extent of polyadenylation [i.e., the average length of the poly(A) tails] in the specific assay but at the same time reduced nonspecific

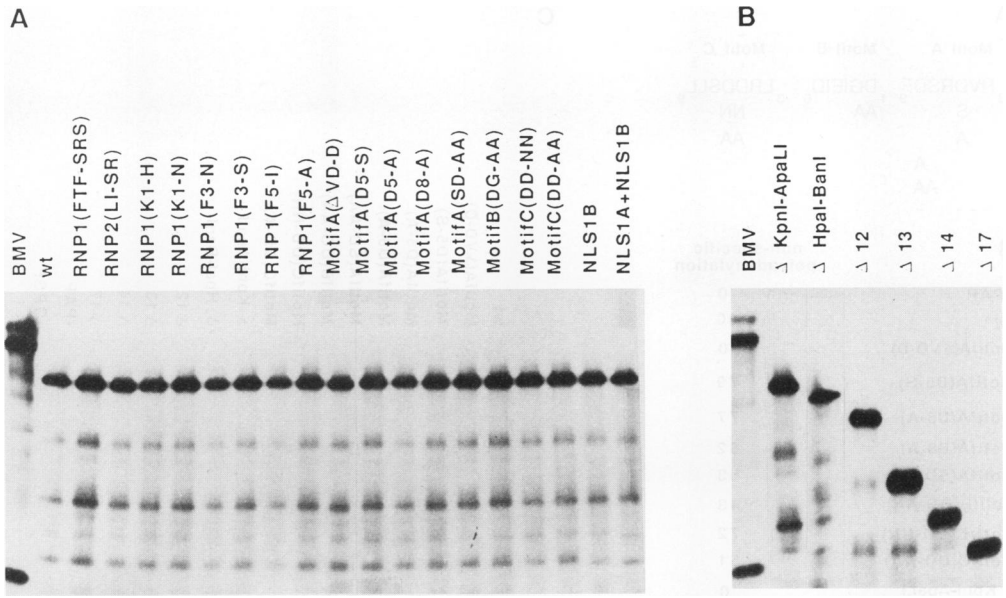

FIG. 2. In vitro-translated PAP derivatives. The indicated mutations were generated in the PAP I cDNA as described in Materials and Methods, using the oligonucleotides listed in Table 1. Mutant and wild-type proteins were prepared by in vitro transcription and translation in a reticulocyte lysate as described in reference 48. [^{3}H]leucine-labeled proteins were resolved on SDS–10% polyacrylamide gels. BMV, protein size standards of ~90 and 30 kDa produced by in vitro translation of brome mosaic virus RNA. (A) Substitution mutants. wt, wild-type PAP I. Mutants are designated by the residues changed in the indicated motifs (see Table 1 and Fig. 1). In all figures, the arrows denoting mutated residues used in the text are replaced with dashes. NLS1A+NLS1B is a double mutant containing the NLS1A and NLS1B mutations. (B) Deletion mutants. The endpoints of the deletions are as described in the text. The first two mutants after the BMV marker are internal deletions, while the remaining mutants are C-terminal truncations.

poly(A) synthesis to 30% of the wild-type level. This is the only mutation among those tested that resulted in enhanced polyadenylation, and possible explanations for this behavior are discussed below. A change of the lysine in position 1 of the octamer to histidine (K-1→H) had little effect on either specific or nonspecific polyadenylation. However, the substitution K-1→N, which eliminates the basic character of this residue, reduced both activities. To obtain evidence that the RNP2 homology was significant, we also tested the effect of a mutation in this motif. The substitution LI→SR in RNP2 reduced nonspecific poly(A) synthesis to 42% of wild-type activity and also slightly reduced specific polyadenylation. Taken together, these results support the idea that PAP contains a functional RBD that is important for both specific and nonspecific polyadenylation. Evidence that these mutations specifically affect binding of the RNA primer is described below.

Certain conserved residues within the putative PM and also sequences outside this region are essential for polyadenylation. We next examined the importance of conserved residues in the putative PM by constructing and analyzing PAP derivatives containing mutations in motifs A, B, and C (Fig. 4). First, we simultaneously deleted the aspartic acid residues in positions 5 and 8 as well as a conserved valine in position 4 of motif A. The resulting triple mutant (ΔVD-D) was completely inactive in both specific and nonspecific polyadenylation. However, when the single aspartic acid residue at position 5 was changed to serine (D-5→S) or alanine (D-5→A), nonspecific and specific activities were both only slightly reduced, indicating that this residue by itself is not essential for catalysis. The aspartic acid at position 8 of motif A is also not essential, since its substitution by alanine (D-8→A) had only a modest effect and a double mutant encompassing this position (SD→AA)

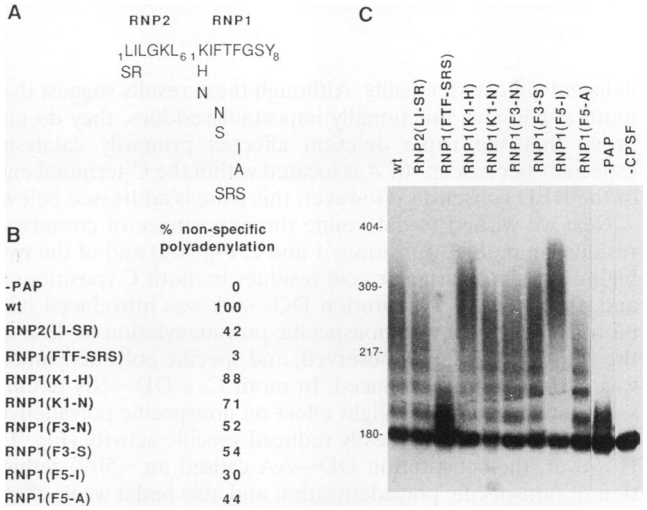

FIG. 3. Mutations in RNP1 and RNP2 inhibit polyadenylation. The mutations within RNP1 and RNP2 elements of the putative PAP RBD (A) and their effects on nonspecific (B) and specific (C) polyadenylation are shown. (A) Letters below the amino acid sequences indicate the introduced residues. SR is a double substitution, and SRS is a triple substitution. Numbers refer to the positions of residues in each motif. (B) Nonspecific poly(A) synthesis activities of the in vitro-translated PAP mutants indicated in panel A were determined by using an AAAAA primer in the presence of 0.5 mM $MnCl_2$ and compared with wild-type PAP I (wt). Numbers are averages of at least three independent experiments. (C) For AAUAAA-dependent polyadenylation, ^{32}P-labeled simian virus 40 late pre-mRNA was incubated with CPSF and the indicated in vitro-translated PAP derivatives as described in Materials and Methods. Processed RNAs were resolved on an 8 M urea gel and autoradiographed. *Msp*I-digested pBR322 DNA was used as a size marker. Numbers indicate sizes in nucleotides.

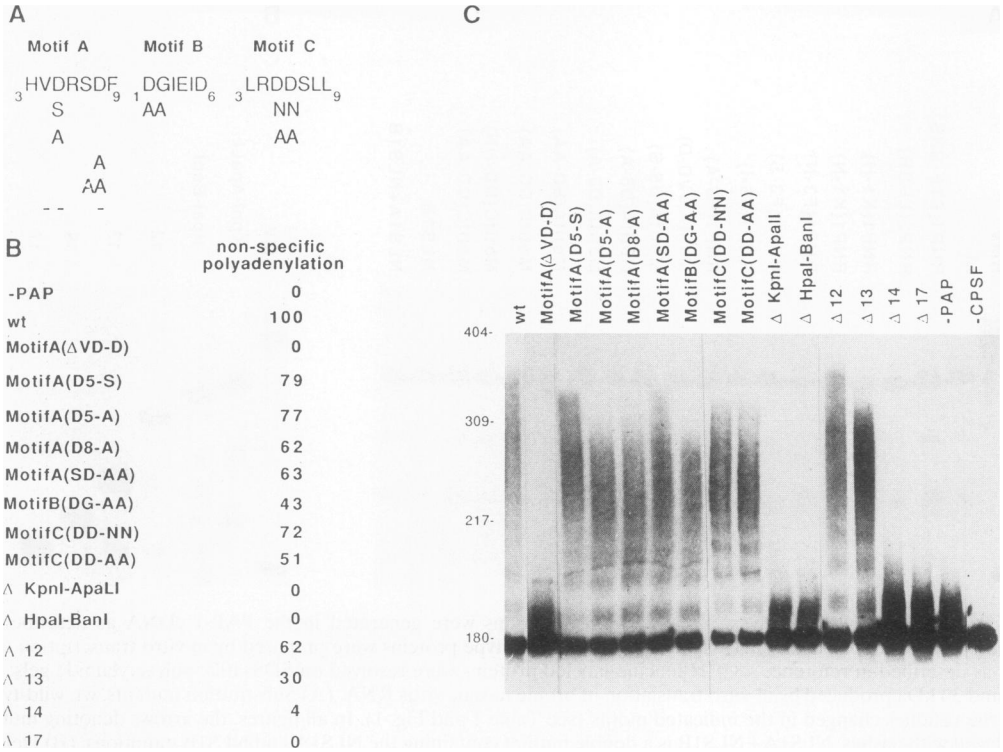

FIG. 4. Mutations in the conserved motifs in the putative PM have variable effects on polyadenylation. (A) Mutations in motifs A to C of the PM are shown. The triple deletion in motif A (ΔVD-D) is symbolized by three minus signs. Nonspecific (B) and specific (C) polyadenylation assays of the substitution mutant PAPs indicated in panel A, as well as the indicated deletion derivatives (see Fig. 2 and text), were carried out exactly as described for Fig. 3. wt, wild type. Sizes are indicated in nucleotides.

behaved almost identically. Although these results suggest that motif A contains functionally important residues, they do not prove that the triple deletion affected primarily catalysis, especially because motif A is located within the C-terminal end of the RBD consensus. However, this issue is addressed below.

Next we wished to determine the importance of conserved residues in motif B (positions 1 and 2; Fig. 4A) and of the two highly conserved aspartic acid residues in motif C (positions 5 and 6). When the substitution DG→AA was introduced into motif B, a reduction of nonspecific polyadenylation to 43% of the wild-type level was observed, and specific polyadenylation was correspondingly reduced. In motif C, a DD→NN double substitution had only a slight effect on nonspecific polyadenylation and also only modestly reduced specific activity (Fig. 4). However, the substitution DD→AA caused an ~50% reduction of nonspecific polyadenylation and also had a weak effect on specific PAP activity, suggesting that these residues play a role, albeit not an essential one, in polyadenylation. Taken together, though, these results argue against the presence of a classical PM in PAP, as mutations of the motif C "double D's" in several RNA-dependent polymerases have been shown to completely inhibit catalysis (see the introduction). However, the presence in PAP of sequence motifs with similarity to motifs A to D and the functional importance of several conserved residues suggest at least some similarity between this region of PAP and a PM.

To determine whether sequences outside the RBD-PM region are necessary for polyadenylation, amino acid residues between 222 and 310 (Δ*Hpa*I-*Ban*I) or 310 and 380 (Δ*Kpn*I-*Apa*LI) (the PM similarity ends at about residue 228) were deleted, and a series of C-terminal truncations, deleting up to ~350 residues, was also generated. In vitro-translated mutant PAPs were again synthesized and used in the two standard polyadenylation assays (Fig. 4). The two adjacent internal deletions were both completely inactive in nonspecific and specific assays. In contrast, C-terminal truncations of 128 (Δ12) or 174 (Δ13) residues all retained significant levels of nonspecific and specific activities (see also reference 48). A sharp drop in both activities was detected between truncations Δ13 and Δ14, which deleted 212 residues. Δ14 retained ~4% nonspecific polyadenylation activity and little, if any, specific activity, and further deletions (e.g., Δ17, which removed 345 residues) generated proteins that were completely inactive. Perhaps significantly, the 38-amino-acid region spanned by the endpoints of Δ13 and Δ14 contains the consensus bipartite NLS (see below).

High primer concentrations can rescue RBD mutations but not mutations in the putative catalytic domain. In the preceding paragraphs, we presented evidence for the presence of an RNP-type RBD and a catalytic region in PAP. But the data presented did not show directly that the so-called RBD mutants were defective because they had a decreased ability to bind the RNA primer, nor did they show that the putative PM mutants were defective for some other reason. Our efforts to measure RNA binding directly have been only partially successful (data not shown), at least in part because of the low affinity of PAP for RNA (65). We therefore developed another assay to address this issue.

Specifically, if a given mutant was defective only because it was unable to bind the RNA primer, then it should be possible to rescue the defect, at least partially, by increasing the primer concentration. In contrast, mutants defective in another step,

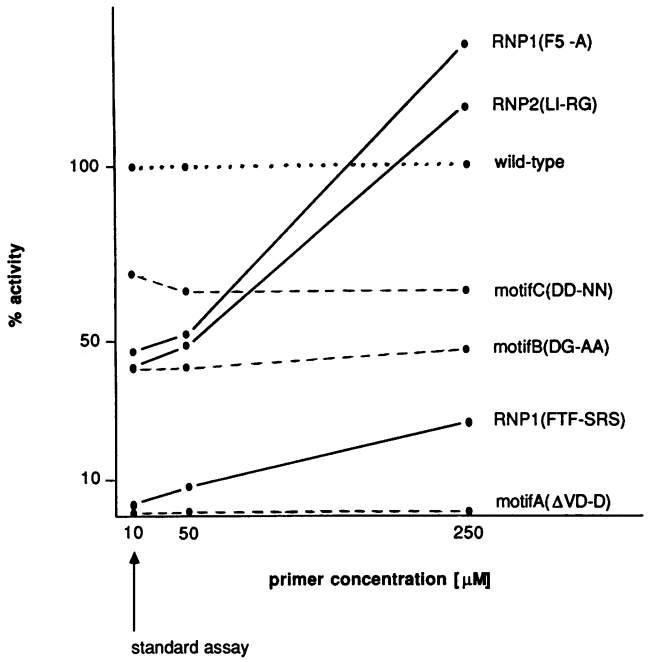

FIG. 5. Dependence of poly(A) synthesis activity on RNA primer concentration. The activities of the indicated PAP I mutant derivatives were measured in nonspecific poly(A) synthesis assays and compared with the activity of wild-type PAP I exactly as in Fig. 3 and 4 except that the oligo(A_5) primer concentration in the reaction mixtures was 10 (standard), 50, or 250 μM. Numbers are averages of at least three independent experiments. Solid lines, RBD mutants; dashed lines, PM mutants; dotted line, wild-type PAP.

e.g., catalysis or ATP binding, should not be rescued at high primer concentrations. Figure 5 shows the effects of increasing the concentration of the oligo(A_5) primer on nonspecific polyadenylation, relative to wild-type PAP, of several derivatives containing mutations within the RBD and PM motifs. Results with three RBD and three PM mutants are shown. The activity of the PAP mutant containing the FTF→SRS substitution in RNP1, which, as shown also in Fig. 3, was nearly inactive at 10 μM primer (the standard concentration) and inactive at 2 μM (results not shown), was increased to almost 25% of the wild-type activity at 250 μM primer. PAP derivatives containing substitutions in RNP1 (F-5→A) and in RNP2 (LI→SR) were both increased from just over 40% of the wild-type activity at 10 μM to levels that were in fact slightly higher than the wild-type level at 250 μM. In sharp contrast, the activities of the PM mutants were not significantly affected by elevated primer concentrations. The derivative containing the ΔVD-D mutation in motif A of the PM was essentially inactive at all primer concentrations tested. Likewise, the enzymes containing the DG→AA mutation in motif B or the DD→NN mutation in motif C, which were both partially active at the standard primer concentration (see also Fig. 4), showed no significant changes in activity at elevated concentrations.

We also tested four additional RBD mutants, four additional PM mutants, and the two large internal deletion mutants (Fig. 3 and 4) at various primer concentrations (results not shown). Three of the additional RBD mutants (K-1→H, K-1→N, and F-3→N) showed significant increases in relative PAP activity, nearly indistinguishable from those observed with the three RBD mutants analyzed in Fig. 5A. One (F-5→I) increased to a lesser degree (from ~30% to only 40% of the wild-type level). While we do not know why this mutant behaved differently from the others, it is perhaps relevant that F-5→I was the only mutant to show elevated activity in the specific polyadenylation assay (Fig. 3C). The four additional PM mutants tested (motif A D-5→A, D-5→S, and D-8→A; motif C DD→AA) behaved essentially identically to the PM mutants shown, none giving a significant increase in activity. In addition, the Δ*Hpa*I-*Ban*I and Δ*Kpn*I-*Apa*LI internal deletions behaved identically to the ΔVD-D deletion, displaying essentially no activity at any primer concentration tested. Thus, six of seven RBD mutants, but none of nine PM mutants (including the two large deletions), showed significant primer-dependent increases in PAP activity. These findings support the hypothesis that the putative RBD is indeed an RNP-type RBD that functions in primer binding and that the RBD and catalytic domains can be experimentally separated from each other.

The exclusively nuclear localization of PAP is mediated by two widely separated bipartite NLS sequences. The role of PAP in AAUAAA-dependent 3′-end processing of pre-mRNA and the presence of a consensus bipartite NLS in PAP I and II (48) suggest that PAP is nuclear. However, this has not been proven, and as discussed in the introduction, other data suggest the existence of PAP activity in the cytoplasm. To investigate the subcellular localization of PAP, we constructed expression vectors that encode PAP derivatives (indicated in Fig. 6A and B) with an N-terminal epitope tag containing nine residues of the influenza virus hemagglutinin antigen (the so-called flu epitope [21]). The cDNAs were cloned downstream of the cytomegalovirus promoter (11) and transfected into human 293 cells. Subcellular localization of the exogenously expressed proteins was determined by indirect immunofluorescence with monoclonal antibody 12CA5, which specifically recognized the flu epitope (see Materials and Methods). Epitope-tagged PAP I and II were both found to accumulate exclusively in nuclei (Fig. 6C, panels 1 and 2). The distribution was smooth and granular, and nucleoli showed markedly less staining.

To begin to examine the sequences important for PAP nuclear localization, we first examined the subcellular localization of several C-terminal deletion mutants of PAP I (Fig. 6A). Unexpectedly, two of these, Δ*Pvu*II (Fig. 6C, panel 3), in which only a fraction of the S/T-rich region was deleted, and Δ13 (Fig. 6C, panel 4), in which almost all of it was deleted (both retained the consensus NLS, NLS1), were localized throughout the cell, indicating that the S/T-rich region contains sequences required for efficient nuclear localization. (The Δ13 derivative also resulted in less intense staining, suggesting that the deleted residues also contain sequences required for maximal accumulation; see below.) However, a C-terminal truncation (Δ14) that was also deleted of NLS1 was exclusively cytoplasmic (Fig. 6C, panel 5), suggesting that sequences encompassing NLS1 can be important for efficient nuclear localization. Consistent with these results, PAP III, the C-terminally truncated form, was largely cytoplasmic (Fig. 6C, panel 6), although significant nuclear accumulation could also be detected, possibly because of PAP III's low molecular mass (~43 kDa), which might allow passive diffusion into the nucleus.

To investigate the role of specific residues in nuclear targeting, we constructed and analyzed a series of substitution mutants (Fig. 6B). Surprisingly, mutation of either or both of the two basic clusters, KRK and KKKK, in NLS1 resulted in decreased protein accumulation but only slightly affected nuclear localization (Fig. 6C, panel 7, and results not shown). These mutants showed a low level of cytoplasmic accumulation, although this is not apparent from the figure because of the overall reduced staining relative to the wild type (compare

panels 1 and 7). These results are consistent with the presence of an additional NLS in PAP. Indeed, we found that a basic motif in the S/T-rich domain, the sequence KKTK (NLS2), is essential for efficient nuclear targeting. This sequence is located 147 residues C terminal to NLS1, just N terminal to the point of divergence between PAP I and PAP II (Fig. 1). The substitution KKTK→NITK resulted in essentially uniform accumulation of PAP throughout the cell (Fig. 6C, panel 8), similar to the C-terminal truncations that retained NLS1 but were deleted of NLS2 (Δ*Pvu*II and Δ13). These results suggest that PAP contains two widely separated NLSs, neither of which is able by itself to direct PAP exclusively to the nucleus. Consistent with this, simultaneous substitutions in both NLS1 and NLS2 caused almost complete cytoplasmic localization (Fig. 6C, panel 9), confirming that both of the NLS sequences are necessary to bring about efficient nuclear localization.

Mutations in NLS1 inhibit specific but not nonspecific polyadenylation. Because the results obtained with the deletion mutants encompassing NLS1 suggested that this region was important for activity in vitro (Fig. 4), we tested two NLS1 substitution mutants in specific and nonspecific polyadenylation assays (Fig. 7). The nonspecific polyadenylation activities of the NLS1A mutant (KKKK→KNNI) and the double mutant NLS1A+NLS1B (KRK→NLN; KKKK→KNNI) were similar to that of wild-type PAP (Fig. 7B). However, specific polyadenylation was reduced by mutation of a single basic cluster and completely inhibited by simultaneous substitution of both basic clusters (Fig. 7C). These results suggest that the basic clusters are not involved in catalysis but function in AAUAAA-dependent polyadenylation, perhaps by facilitating interaction with CPSF (see the introduction).

To investigate these unexpected findings further, we tested the effects of adding basic or neutral polyamino acids to specific and nonspecific polyadenylation assays (Fig. 8). If the inhibitory effects of the NLS1 mutations reflected interaction with another factor and if basicity were sufficient for this, then basic polyamino acids might inhibit specific but not nonspecific polyadenylation in *trans*. These assays used PAP extensively purified from HeLa cells (53a) and purified bovine CPSF. Strikingly, addition of neutralized polylysine or polyarginine, but not polyalanine or polyasparagine, to 10 μM completely inhibited specific but not nonspecific polyadenylation. In fact, the basic polyamino acids actually stimulated nonspecific polyadenylation, by 1.3-fold (polylysine) and 2.6-fold (polyarginine). Together, these results suggest an important role for the NLS1 basic clusters in AAUAAA-dependent polyadenylation.

The S/T-rich domain is phosphorylated in vivo. To investigate whether the C-terminal S/T-rich region is subject to posttranslational modification, we examined flu epitope-tagged wild-type and mutant PAP derivatives expressed in transfected cells by Western blotting (see Materials and Methods). Analysis of full-length PAP I revealed a band corresponding to a protein of the expected size, 77 kDa, as well as multiple species in the size range of 80 to 100 kDa (Fig. 9A). In contrast, the deletion mutant lacking the S/T-rich region, Δ13, gave rise to a single species of the predicted size (59 kDa; Fig. 9A). The level of expression for mutants lacking the entire S/T-rich region (e.g., Δ13) was significantly lower than observed with wild-type or mutants retaining at least part of the S/T-rich region (e.g., Δ*Pvu*II), which is consistent with the levels of expression detected by indirect immunofluorescence microscopy (Fig. 6C and results not shown). These findings indicate that the S/T-rich region is heavily modified in transfected cells and is also required for maximal protein accumulation. To determine whether some or all of this modification resulted from phosphorylation, extract from cells transfected with full-length PAP I was incubated with potato acid phosphatase. Essentially only two major bands, of about 77 and 84 kDa, were detected following phosphatase treatment (Fig. 9B). Therefore, most if not all of the high-molecular-weight (high-MW) forms of PAP appear to arise from phosphorylation. Since two of the NLS substitution mutations described above, NLS1A (KKKK→KNNI) and NLS2 (KKTK→NITK), also destroyed putative kinase sites, namely, a cyclic AMP (cAMP)-dependent kinase site and a cdc2 kinase site, respectively, the high-MW pattern of these substitution mutants was also examined (Fig. 9A). The NLS1 mutation inhibited expression of some of the high-MW forms, and generally weaker expression of the remaining bands was reproducibly observed, consistent with the results obtained by indirect immunofluorescence. Strikingly, the NLS2 substitution caused the disappearance of all high-MW forms, resulting in a pattern similar to that of phosphatase-treated PAP. These results suggest that both of the NLS sites are phosphorylated and that phosphorylation of the cdc2 site may be critical for additional modifications.

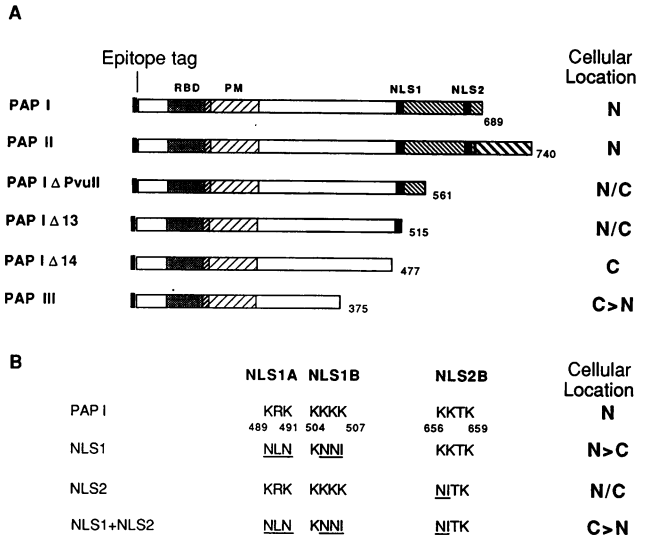

FIG. 6. NLS1 and NLS2 both function to localize PAP to the nucleus. Wild-type and mutant epitope-tagged PAP constructs were transiently transfected into 293 cells, and cellular location was determined by indirect immunofluorescence microscopy using an anti-epitope antibody as described in Materials and Methods. (A) Diagrams of PAP derivatives analyzed. Features of the proteins are as indicated in Fig. 1. PAP I and II, full-length bovine PAPs; PAP III, truncated naturally occurring human derivative of PAP; PAP I Δ*Pvu*II, Δ13, and Δ14, C-terminal truncations of PAP I. (B) PAP I mutants containing the indicated amino acid substitutions in three C-terminal basic clusters. Mutated residues are underlined. N, exclusively nuclear staining; N>C, nuclear, some cytoplasmic staining; N/C, equally strong staining in nucleus and cytoplasm; C>N, cytoplasmic, some nuclear staining; C, exclusively cytoplasmic staining. Numbers indicated C-terminal residues of the indicated PAP derivative (A) or first and last residues of the basic clusters (B). (C) Indirect immunofluorescence microscopy of cells transfected with the wild-type and mutant PAPs shown in panels A and B. 1, PAP I; 2, PAP II; 3, PAP I Δ*Pvu*II; 4, PAP I Δ13; 5, PAP I Δ14; 6, PAP III; 7, PAP I NLS1; 8, PAP I NLS2; 9, PAP I NLS1+NLS2.

DISCUSSION

The results described above present insights into the structure and function of PAP. We have defined several functional

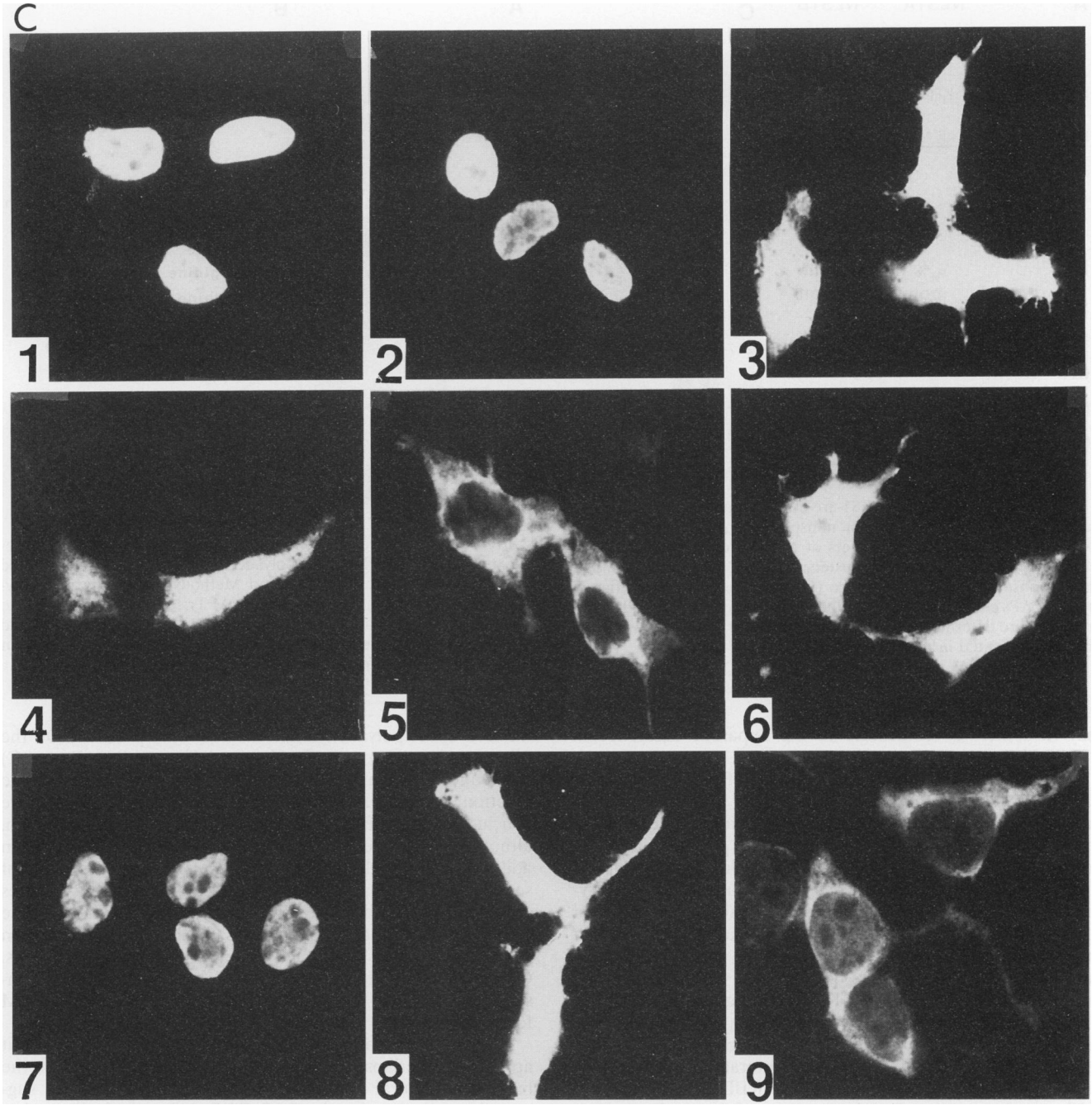

FIG. 6—*Continued.*

domains and provided evidence that two of these overlap in novel ways. Specifically, at the N terminus, a classical RNP-type RBD appears to overlap a region essential for catalysis, while one of the two C-terminal signals important for nuclear localization is also required for AAUAAA-dependent polyadenylation. Below we discuss the significance of each of these functional domains.

The RBD. Our data provide considerable support for the notion that PAP contains a classical RNP-type RBD. Neither full-length PAP (65) nor truncated derivatives lacking C-terminal sequences bind RNA efficiently in vitro (unpublished data). However, on the basis of effects on PAP activity of mutations targeted to residues predicted to affect RNA binding and our observation that these mutations, but not others, can be rescued by elevated primer concentrations, we believe the evidence is strong that the homology we described previously (48) does in fact reflect a bona fide RBD. The presence of an RBD in PAP was somewhat unexpected: no other known polymerase (RNA or DNA) contains such a domain, with a simple explanation being that polymerases by definition must interact only transiently with primer and/or template. PAP is of course distinctive in its lack of a template requirement, and this

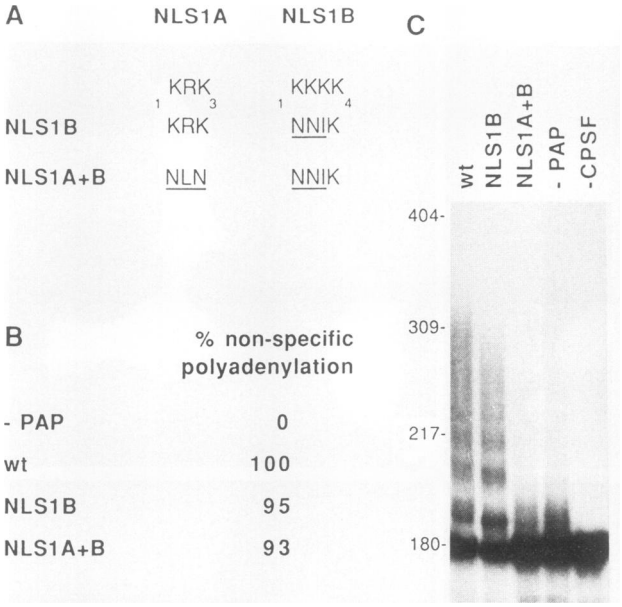

FIG. 7. Basic residues in NLS1 are essential for AAUAAA-dependent polyadenylation but not for nonspecific poly(A) synthesis. (A) Mutations in the two basic clusters of NLS1. Numbers indicate first and last residues of the basic clusters. Mutated residues are underlined. Nonspecific (B) and specific (C) polyadenylation activities of mutant PAPs containing the mutations indicated in panel A and shown in Fig. 2 were determined exactly as described for Fig. 3. wt, wild type. Sizes are indicated in nucleotides.

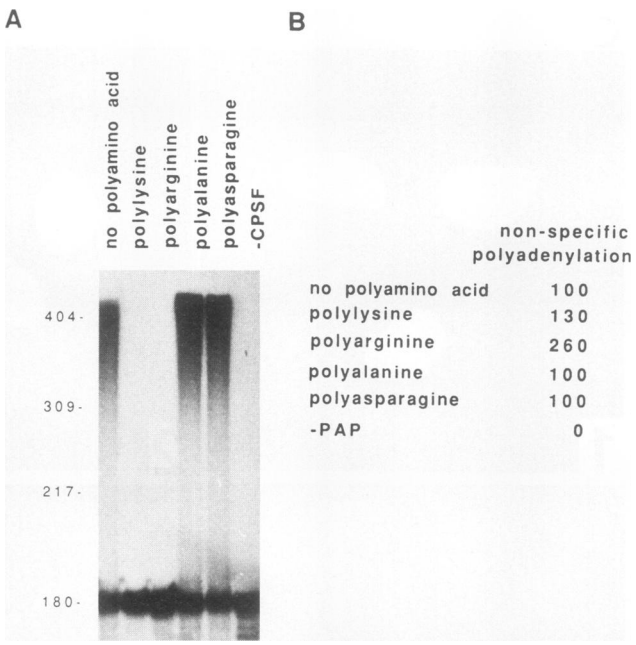

FIG. 8. Basic polyamino acids inhibit specific but enhance nonspecific polyadenylation. Standard polyadenylation assays were performed with HeLa PAP (see Materials and Methods). The indicated polyamino acids were neutralized with 10 mM Tris (pH 8.0) and added to reaction mixtures at a final concentration of 10 μM. Specific (A) and nonspecific (B) polyadenylation assays were otherwise carried out exactly as described for Fig. 3. Sizes are indicated in nucleotides.

may in some way contribute to the apparent difference in structural domains.

We have shown that triple substitution of residues 3 to 5 of RNP1 caused an almost complete loss of specific as well as nonspecific polyadenylation activity, while a number of single substitutions in both RNP1 and RNP2 partially inhibited polyadenylation. Interestingly, several single substitutions caused greater inhibition of nonspecific than of specific polyadenylation, and one (F-5→I) even enhanced the level of specific polyadenylation. A possible explanation for this is that in specific polyadenylation, PAP binds the pre-mRNA cooperatively with CPSF, which is bound to the AAUAAA signal, rather than through protein-RNA contacts alone (44). In contrast, for nonspecific poly(A) synthesis, the PAP RBD is probably solely responsible for recognition of the RNA 3′ end. Thus, it may be that protein-protein interactions can at least partially compensate for defects in the RBD.

The catalytic region. For the purposes of this study, we have assumed that mutants defective in nonspecific polyadenylation that cannot be rescued by elevated primer concentrations, and which lie N terminal to NLS1 (see below), are defective in catalysis. On the basis of distribution of key conserved residues, we had previously suggested (48) that PAP contains a PM similar to that found in several RNA-dependent polymerases (46). Here we have shown that targeted mutations of key residues in the putative PM have variable effects on PAP activity. While these results support the view that the PAP catalytic domain encompasses the previously postulated region, it appears distinct from the prototypical PM domain. Most telling were mutations of the highly conserved double D's in motif C. As mentioned above, substitutions of these two residues in three RNA-dependent polymerases completely inhibit polymerization, while similar substitutions in the PAP motif C affect polyadenylation at most only partially. On the other hand, mutations in the three motifs tested (A, B, and C) did reduce polyadenylation activity, and a small three-residue deletion in motif A was completely inactive. We therefore conclude that the PAP PM is a new type of catalytic domain, although it may be related to previously described PMs. In addition, our results show that an ~270-residue stretch of sequence C terminal to the PM region contains residues essential for specific and nonspecific polyadenylation. These residues might function in catalysis, as deletion mutations in this region cannot be rescued by high primer concentrations. Other RNA-dependent polymerases (e.g., HIV reverse transcriptase [12]) also appear to require residues outside the PM for catalysis.

An intriguing feature of the PAP catalytic domain is that it apparently overlaps with the RBD. Specifically, largely on the basis of structural predictions of Kenan et al. (28), we suggested a structure for the PAP RBD (48), which, like other RBDs, consists of a four-stranded β sheet supported by two α helices. Motif A of the PM was predicted to be largely in a loop separating the second α helix and fourth β strand. The results described here show that residues in this region are important for polyadenylation and that they appear to function in catalysis, since none of the PAP derivatives containing motif A mutations could be rescued by elevated primer concentrations. There are at least two other examples in which residues within an RBD play roles distinct from RNA binding. In the U2 small nuclear RNP B″ protein, residues in one RBD are essential for interaction with another U2 protein, A′ (8, 56). In the splicing factor ASF/SF2, RBD mutations that have no detectable effect on RNA binding but which prevent the protein from functioning in splicing have been described (73). Finally, it is noteworthy that the structural arrangement we suggest for the PAP

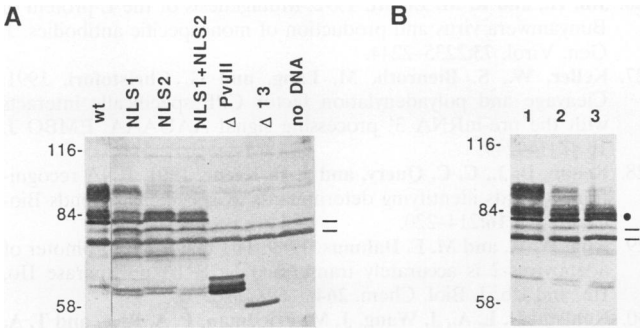

FIG. 9. The S/T-rich region is extensively phosphorylated in vivo. 293 cells were transiently transfected with wild-type or mutant flu epitope-tagged PAP cDNA constructs as described for Fig. 6. Cell lysates were prepared and subjected to SDS-PAGE. Western blots were done as described in Materials and Methods, using the anti-flu epitope monoclonal antibody 12CA5. Two cross-reacting bands at ~72 and ~68 kDa (horizontal bars) are present in all lanes. Numbers indicate sizes (in kilodaltons) of standard proteins. (A) wt, wild-type PAP I. The indicated mutants are as shown in Fig. 6A and B. No DNA, extract from mock-transfected cells. (B) Ten microliters of whole-cell extract from cells transfected with the flu epitope-tagged wild-type PAP I construct was incubated with 50 mM phosphatase inhibitor NaF for 30 min at 37°C (lane 1) or with 1 (lane 2) and 4 (lane 3) μl of potato acid phosphatase for 30 min at 37°C. The dot indicates the position of the two major PAP bands of ~77 and ~84 kDa that remain after phosphatase treatment.

RBD and PM may be similar to the corresponding regions of HIV reverse transcriptase (30), which not only are in the same linear order in the protein sequence but also appear to overlap in a manner similar to that suggested here for PAP.

The NLSs. We have shown that PAP I and PAP II expressed in transiently transfected cells are exclusively nuclear in the vast majority of transfected cells. The uniform, granular staining observed suggests that PAP is distributed relatively uniformly throughout the nucleus. While it is conceivable that the overexpression that likely occurs during transient transfection obscures a more specific staining pattern, this pattern is similar to that observed endogenously with another polyadenylation factor, CstF (58). It is interesting to compare this distribution with that observed with many splicing factors, expressed endogenously or exogenously, which are concentrated in discrete speckles or foci (14, 22, 34, 57). One interpretation of these observations is that polyadenylation occurs throughout the nucleus, perhaps on nascent pre-mRNAs, while splicing occurs in distinct subnuclear structures.

PAP requires the function of two NLSs, separated by ~140 residues, to target it efficiently to the nucleus. Each of these, in turn, appears to be a bipartite signal in which two basic clusters are separated by ~10 residues (although we have not tested the requirement of the two upstream basic residues in the second NLS). Somewhat surprisingly, because the upstream signal is the strongest match to the bipartite NLS consensus (18), the second NLS appears stronger, as its substitution leads to significantly greater cytoplasmic accumulation than does mutation of both basic clusters in NLS1. Although several other nuclear proteins have been shown to require more than one NLS (43, 70, 72), these are all clustered within ~40 residues, and PAP is thus unique in requiring two widely separated signals for nuclear transport. The unexpected finding that NLS1 is essential for specific polyadenylation shows to our knowledge for the first time that an NLS can have two apparently unrelated functions. It is conceivable, however, that both functions are mediated by similar mechanisms, namely, protein-protein interactions involving the basic clusters. Basic clusters that function as part of an NLS have been shown to bind specifically to a number of proteins (39). CPSF, the only other factor in addition to PAP essential for AAUAAA-dependent polyadenylation in vitro, might bind to a region of PAP encompassing NLS1, thereby facilitating interaction of PAP with the pre-mRNA. NLS1 overlaps a putative cAMP-dependent protein kinase site, phosphorylation of which could regulate one or both of its functions. Phosphorylation of overlapping kinase sites has been shown to regulate the activity of NLS sequences in other proteins (41, 49).

PAP III is a third form of PAP that also appears to arise from alternative splicing and is truncated well N terminal to NLS1. Consistent with our mutational analysis, PAP III is totally inactive in in vitro assays, and its function is thus enigmatic. It seems not to be the product of a cloning artifact, as we (unpublished results) and Wahle et al. (67) independently isolated nearly identical cDNAs from HeLa libraries. Also in keeping with mutational studies, a large fraction of PAP III produced in transfected cells is cytoplasmic. This raises the intriguing possibility that PAP III functions in some aspect of poly(A) metabolism in the cytoplasm. Several studies indeed suggest that poly(A) tails can be lengthened as well as shortened in the cytoplasm (reviewed in references 4, 9, and 25). One possibility is that PAP III requires an additional, unknown cytoplasmic factor(s) for activity.

The S/T-rich domain. We have shown that the S/T-rich domain of PAP is extensively phosphorylated in vivo. This region enhances but is not essential for polyadenylation in vitro, and the degree of stimulation increases with its length. An interesting analogy to the S/T-rich region might be the C-terminal domain (CTD) of the RNA polymerase II large subunit, which consists of a heptapeptide repeat that is also S/T rich (2, 17). Like the S/T-region in PAP, the CTD is not essential for transcription in vitro (29). However, the CTD stimulates transcription in a length-dependent manner and influences formation and activation of the transcription complex through its reversible phosphorylation (31, 64). The unphosphorylated form binds the preinitiation complex, and phosphorylation of the CTD appears to facilitate initiation. In a similar way, phosphorylation of the PAP S/T-rich domain might help to regulate formation of the polyadenylation complex and/or the actual initiation of poly(A) synthesis. The fact that phosphorylation sites appear to overlap each NLS also raises the possibility that PAP subcellular localization is subject to control by phosphorylation.

ACKNOWLEDGMENTS

We thank Lai Zhong and Mathew Wang for excellent technical help and the members of the Manley laboratory for helpful advice. We thank Jeff Field for antibody 12CA5 and Ron Prywes for the pCGN vector.

This work was supported by NIH grant 28983.

REFERENCES

1. **Adam, S. A., T. Nakagawa, M. S. Swanson, T. K. Woodruff, and G. Dreyfuss.** 1986. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. Mol. Cell. Biol. **6**:2932–2943.
2. **Allison, L. A., M. Moyle, M. Shales, and C. J. Ingles.** 1985. Extensive homology among the largest subunits of eucaryotic RNA polymerases. Cell **42**:599–610.
3. **Argos, P.** 1988. A sequence motif in many polymerases. Nucleic Acids Res. **16**:9909–9916.
4. **Bachvarova, F.** 1992. A maternal tail of poly(A): the long and the short of it. Cell **69**:895–897.

5. **Bandziulis, R. J., M. S. Swanson, and G. Dreyfuss.** 1989. RNA binding proteins as developmental regulators. Genes Dev. **3**:431–437.
6. **Bardwell, V. J., M. Wickens, S. Bienroth, W. Keller, B. Sproat, and A. I. Lamond.** 1991. Site-directed ribose methylation identifies 2′OH groups in polyadenylation substrates critical for AAUAAA recognition and poly(A) addition. Cell **65**:125–133.
7. **Bardwell, V. J., D. Zarkower, M. Edmonds, and M. Wickens.** 1990. The enzyme that adds poly(A) to mRNAs is a classical poly(A) polymerase. Mol. Cell. Biol. **10**:846–849.
8. **Bentley, R. C., and J. D. Keene.** 1991. Recognition of U1 and U2 small nuclear RNAs can be altered by a five-amino-acid segment in the U2 small nuclear ribonucleoprotein particle (snRNP) B″ protein and through interactions with U2 snRNP-A′ protein. Mol. Cell. Biol. **11**:1829–1839.
9. **Bernstein, B., and J. Ross.** 1989. Poly(A), poly(A) binding protein and the regulation of mRNA stability. Trends Biochem. Sci. **14**:373–377.
10. **Bienroth, S., E. Wahle, E. Satler, C. Crazzarola, and W. Keller.** 1991. Purification of the cleavage and polyadenylation factor involved in 3′ processing of mRNA precursors. J. Biol. Chem. **266**:19768–19776.
11. **Boshart, M., F. Weber, G. Jahn, K. Dorsch-Haesler, B. Fleckenstein, and W. Schaffner.** 1985. A very strong enhancer located upstream of the immediate early gene of human cytomegalovirus. Cell **41**:521–530.
12. **Boyer, P. L., A. L. Ferris, and S. H. Hughes.** 1992. Cassette mutagenesis of the reverse transcriptase of human immunodeficiency virus type 1. J. Virol. **66**:1031–1039.
13. **Brennan, C. A., and T. Platt.** 1991. Mutations in an RNP1 consensus sequence of rho protein reduce RNA binding affinity but facilitate helicase turnover. J. Biol. Chem. **266**:17269–17305.
14. **Carmo-Fonseca, M., R. Pepperkok, B. S. Sproat, W. Ansorge, M. S. Swanson, and A. I. Lamond.** 1991. In vivo detection of snRNP-rich organelles in the nuclei of mammalian cells. EMBO J. **18**:1863–1873.
15. **Christofori, G., and W. Keller.** 1988. 3′-cleavage and polyadenylation of mRNA precursors *in vitro* requires a poly(A) polymerase, a cleavage factor, and a snRNP. Cell **54**:875–889.
16. **Christofori, G., and W. Keller.** 1989. Poly(A) polymerase purified from HeLa cell nuclear extract is required for both cleavage and polyadenylation of pre-mRNA in vitro. Mol. Cell. Biol. **9**:193–203.
17. **Corden, J. L., D. L. Ladena, J. M. Ahearn, and D. E. Dahmus.** 1985. A unique structure at the carboxyl terminus of the largest subunit of eucaryotic RNA polymerase II. Proc. Natl. Acad. Sci. USA **82**:7934–7938.
18. **Dingwall, C., and R. L. Laskey.** 1991. Nuclear targeting sequences—a consensus? Trends Biochem. Sci. **16**:478–481.
19. **Dreyfuss, G., M. J. Matunis, S. Piñol Roma, and C. G. Burd.** 1993. hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochem. **62**:289–321.
20. **Edmonds, M.** 1982. Poly(A) adding enzymes, p. 217–244. *In* P. D. Boyer (ed.), The enzymes, vol. XV. Academic Press, New York.
21. **Field, J., J.-I. Nikama, D. Brock, C. C. McDonald, L. Rodgen, I. Wilson, A. Lerner, and M. Wigler.** 1988. Purification of a RAS-responsive adenylyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method. Mol. Cell. Biol. **8**:2159–2165.
22. **Fu, X.-D., and T. Maniatis.** 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. Nature (London) **343**:437–441.
23. **Gilmartin, G. M., and J. R. Nevins.** 1989. An ordered pathway of assembly of components required for polyadenylation site recognition and processing. Genes Dev. **3**:2180–2189.
24. **Hoffman, D. W., C. C. Query, B. Golden, S. W. White, and J. D. Keene.** 1991. RNA-binding domain of the A protein component of the U1 small ribonucleoprotein analyzed by NMR spectroscopy is structurally similar to ribosomal proteins. Proc. Natl. Acad. Sci. USA **88**:2495–2499.
25. **Jackson, R. J., and N. Standart.** 1990. Do the poly(A) tail and the 3′ untranslated region control mRNA translation? Cell **62**:15–24.
26. **Jin, H., and R. M. Elliott.** 1992. Mutagenesis of the L protein of Bunyamwera virus and production of monospecific antibodies. J. Gen. Virol. **73**:2235–2244.
27. **Keller, W., S. Bienroth, M. Lang, and G. Christofori.** 1991. Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3′ processing signal AAUAAA. EMBO J. **10**:4241–4249.
28. **Kenan, D. J., C. C. Query, and J. D. Keene.** 1991. RNA recognition: towards identifying determinants of specificity. Trends Biochem. Sci. **16**:214–220.
29. **Kim, W. Y., and M. E. Dahmus.** 1989. The major late promoter of adenovirus-2 is accurately transcribed by RNA polymerase IIo, IIa, and IIb. J. Biol. Chem. **264**:3169–3176.
30. **Kohlstaedt, L. A., J. Wang, J. M. Friedman, P. A. Rice, and T. A. Steitz.** 1992. Crystal structure at 3.5 Å resolution of HIV-reverse transcriptase complexed with an inhibitor. Science **256**:1783–1790.
31. **Koleske, A. J., S. Buratowski, M. Nonet, and R. A. Young.** 1992. A novel transcription factor reveals a functional link between the RNA polymerase II CTD and TFIID. Cell **69**:883–894.
32. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. Nature (London) **227**:680–685.
33. **Larder, B. A., D. J. M. Purifoy, K. L. Powell, and G. Darby.** 1987. Site-specific mutagenesis of AIDS virus reverse transcriptase. Nature (London) **327**:716–717.
34. **Lerner, E. A., M. R. Lerner, C. A. Janeway, and J. A. Steitz.** 1981. Monoclonal antibodies to nucleic acid containing cellular constituents: probes for molecular biology and autoimmune disease. Proc. Natl. Acad. Sci. USA **78**:2737–2741.
35. **Lingner, J., J. Kellermann, and W. Keller.** 1991. Cloning and expression of the essential gene for poly(A) polymerase from *S. cerevisiae*. Nature (London) **354**:496–498.
36. **Lowe, M., V. Parmar, S. P. Kemp, and B. A. Larder.** 1991. Mutational analysis of two conserved sequence motifs in HIV reverse transcriptase. FEBS Lett. **282**:231–234.
37. **Manley, J. L.** 1983. Accurate and specific polyadenylation of mRNA precursors in a soluble whole-cell lysate. Cell **33**:595–605.
38. **Manley, J. L.** 1988. Polyadenylation of mRNA precursors. Biochim. Biophys. Acta **950**:1–12.
39. **Meier, U. T., and G. Blobel.** 1992. Nopp 140 shuttles on tracks between nucleolus and cytoplasm. Cell **70**:127–38.
40. **Merrill, B. M., K. L. Stone, F. Lobianchi, S. H. Wilson, and K. R. Williams.** 1988. Phenylalanines that are conserved among several RNA-binding proteins form part of a nucleic acid-binding pocket in the A1 heterogeneous nuclear ribonucleoprotein. J. Biol. Chem. **263**:3307–3313.
41. **Moll, T., G. Tebb, U. Surama, H. Robitsch, and K. Nasmyth.** 1991. The role of phosphorylation and the CDC28 protein kinase in cell cycle-regulated nuclear transport of the *S. cerevisiae* transcription factor SW15. Cell **66**:743–758.
42. **Moore, C. L., and P. A. Sharp.** 1985. Accurate cleavage and polyadenylation of exogenous RNA substrate. Cell **41**:845–855.
43. **Morin, N., C. Delsert, and D. F. Klessig.** 1989. Nuclear localization of the adenovirus DNA-binding protein: requirement for two signals and complementation during viral infection. Mol. Cell. Biol. **9**:4372–4380.
44. **Murthy, K. G., and J. L. Manley.** 1992. Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. J. Biol. Chem. **267**:14804–14811.
45. **Nagai, K., C. Oubridge, T. H. Jessen, J. Li, and P. Evans.** 1990. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. Nature (London) **348**:515–520.
46. **Poch, O., I. Sauvaget, M. Delarue, and N. Tordo.** 1989. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. EMBO J. **8**:3867–3874.
47. **Query, C. C., R. C. Bentley, and J. D. Keene.** 1989. A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. Cell **57**:89–101.
48. **Raabe, T., F. J. Bollum, and J. L. Manley.** 1991. Primary structure and expression of bovine poly(A) polymerase. Nature (London) **353**:229–234.

48a. **Raabe, T., and J. L. Manley.** Unpublished data.

49. **Rhis, H.-P., and R. Peters.** 1989. Nuclear transport kinetics

depend on phosphorylation-site-containing sequences flanking the karyophilic signal of the simian virus 40 T-antigen. EMBO J. **8**:1479–1484.

50. **Ribas, J. C., and R. B. Wickner.** 1992. RNA-dependent RNA polymerase consensus sequence of the L-A double-stranded RNA virus: definition of essential domains. Proc. Natl. Acad. Sci. USA **89**:2185–2189.
51. **Richter, J.** 1991. Translational control during early development. Bioessays **13**:179–183.
52. **Robbins, J., S. M. Dilworth, R. A. Laskey, and C. Dingwall.** 1989. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. Cell **64**:615–624.
53. **Ryner, L. C., Y. Takagaki, and J. L. Manley.** 1989. Requirements for accurate and efficient mRNA 3′-end cleavage and polyadenylation of a simian virus 40 early pre-RNA in vitro. Mol. Cell. Biol. **9**:1759–1771.

53a. **Ryner, L. C., Y. Takagaki, and J. L. Manley.** 1989. Multiple forms of poly(A) polymerase purified from HeLa cells function in specific mRNA 3′ end formation. Mol. Cell. Biol. **9**:4229–4238.

54. **Sachs, A. B., M. W. Bond, and R. D. Kornberg.** 1986. A single gene from yeast for both nuclear and cytoplasmic polyadenylate-binding protein: domain structure and expression. Cell **45**:827–835.
55. **Sachs, A. B., R. W. Davis, and R. D. Kornberg.** 1987. A single domain of poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. Mol. Cell. Biol. **7**:3268–3276.
56. **Scherly, D., W. Boelens, N. A. Dathan, W. J. van Venrooij, and I. W. Mattaj.** 1990. Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B″ and their cognate RNAs. Nature (London) **345**:502–506.
57. **Spector, D. L.** 1990. Higher order nuclear organization: three dimensional distribution of small nuclear ribonucleoprotein particles. Proc. Natl. Acad. Sci. USA **87**:147–151.
58. **Takagaki, Y., J. L. Manley, C. C. MacDonald, J. Wilusz, and T. Shenk.** 1990. A multisubunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs. Genes Dev. **4**:2112–2120.
59. **Takagaki, Y., L. C. Ryner, and J. L. Manley.** 1988. Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation. Cell **52**:731–742.
60. **Takagaki, Y., L. C. Ryner, and J. L. Manley.** 1989. Four factors are required for 3′-end cleavage of pre-mRNAs. Genes Dev. **3**:1711–1724.
61. **Tanaka, M., and W. Herr.** 1990. Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. Cell **60**:375–386.
62. **Terns, M. P., and S. T. Jacob.** 1989. Role of poly(A) polymerase in the cleavage and polyadenylation of mRNA precursor. Mol. Cell. Biol. **9**:1435–1444.
63. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
64. **Usheva, A., E. McDonald, A. Goldring, H. Lu, L. Houbaui, D. Reistory, and Y. Atoni.** 1992. Specific interaction between nonphosphorylated form of RNA polymerase II and the TATA binding protein. Cell **69**:871–881.
65. **Wahle, E.** 1991. Purification and characterization of a mammalian poly(A) polymerase involved in the 3′-end processing of messenger RNA precursor. J. Biol. Chem. **266**:3131–3139.
66. **Wahle, E., and W. Keller.** 1992. The biochemistry of 3′-end cleavage and polyadenylation of messenger RNA precursors. Annu. Rev. Biochem. **61**:419–440.
67. **Wahle, E., G. Martin, E. Schiltz, and W. Keller.** 1991. Isolation and expression of cDNA clones encoding mammalian poly(A) polymerase. EMBO J. **10**:4251–4257.
68. **Wickens, M.** 1990. How the messenger got its tail: addition of poly(A) in the nucleus. Trends Biochem. Sci. **15**:277–281.
69. **Wittekind, M., M. Görlach, M. Friedrichs, G. Dreyfuss, and L. Mueller.** 1992. ^{1}H, ^{13}C and ^{15}N NMR assignments and global folding pattern of the RNA-binding domain of the human hnRNP C proteins. Biochemistry **31**:6254–6265.
70. **Ylikomi, T., M. T. Bocquel, M. Berry, H. Gronemeyer, and P. Chambon.** 1992. Cooperation of proto-signals for nuclear accumulation of estrogen and progesterone receptors. EMBO J. **11**:3681–3694.
71. **Zarkower, D., R. Stephenson, M. Sheets, and M. Wickens.** 1986. The AAUAAA sequence is required both for cleavage and for polyadenylation of simian virus 40 pre-mRNA in vitro. Mol. Cell. Biol. **6**:2317–2323.
72. **Zhao, L.-J., and R. Padmanabhan.** 1991. Three basic regions in adenovirus DNA polymerase interact differentially depending on the protein context to function as bipartite nuclear localization signals. New Biol. **3**:1074–1088.
73. **Zuo, P., and J. L. Manley.** 1993. Functional domains of the human splicing factor ASF/SF2. EMBO J. **12**:4727–4737.