The EMBO Journal vol.10 no.1 pp.215-219, 1991

Poly(A) site efficiency reflects the stability of complex formation involving the downstream element

Elizabeth A.Weiss, Gregory M.Gilmartin and Joseph R.Nevins

Howard Hughes Medical Institute, Department of Microbiology and Immunology, Duke University Medical Center, PO Box 3054, Durham, NC 27710, USA

Communicated by L.Philipson

A critical step in mRNA biogenesis is the generation of the mRNA 3′ end through an endonucleolytic cleavage of the primary transcript followed by the addition of a ~200 nucleotide (nt) poly(A) tail. The efficiency of poly(A) site function can vary widely and for those genes with multiple poly(A) sites, the choice can be a regulated event. A functional poly(A) site is characterized by *cis*-acting RNA sequences including the well-conserved AAUAAA hexamer, located 10-30 nt upstream of the cleavage site, and a highly variable downstream GU- or U-rich element. The gene specific nature of the downstream sequence suggests that it may be a primary determinant of poly(A) site efficiency. Several recent studies have detailed the purification of factors that mediate the cleavage and polyadenylation reaction and that recognize the *cis*-acting signals. Two of these factors are responsible for the formation of a stable, committed ternary complex with the pre-RNA. In order to define the role of this stable complex in poly(A) site function, we have compared the processing efficiency of several pre-mRNAs with the stability of the complex that forms on these RNAs. We show that ternary complex stability reflects both the *in vivo* and the *in vitro* efficiency of the poly(A) site and that the stability of this complex is dependent on the nature of the downstream sequence element. We conclude that the stability of these protein-RNA interactions, dictated by the downstream element, plays a major role in determining the processing efficiency of a particular poly(A) site.

Key words: complex stability/polyadenylation efficiency/protein-RNA interaction

Introduction

The biogenesis of a eukaryotic mRNA is a multi-step process that requires a series of RNA processing events (Nevins, 1983). A critical step common to virtually all mRNAs is the generation of the mature 3′ end of the mRNA that is subsequently polyadenylated (Birnstiel *et al.*, 1985). In addition to the fact that this reaction is essential for the formation of a functional mRNA, it also appears that poly(A) site utilization can be subject to variation and regulation. The clearest example of a regulatory role for polyadenylation is found in the differentiating B cell where the switch from a membrane form of the immunoglobulin heavy chain to a secreted form of the polypeptide is governed in part through alternative use of poly(A) sites (Galli *et al.*, 1987).

A final understanding of the mechanisms of poly(A) site formation as well as the basis for regulation of this event will require the identification of the factors in nuclei that carry out the reaction. One presumes that a change in utilization of a particular poly(A) site must be due to changes or alterations in a rate-limiting factor or factors that directs processing dependent on the critical *cis*-acting RNA sequences (Proudfoot and Brownlee, 1976; Fitzgerald and Shenk, 1981; Simonsen and Levinson, 1983; McDevitt *et al.*, 1984; Gil and Proudfoot, 1984; Sadofsky and Alwine, 1984; Woychik *et al.*, 1984; Mclauchlan *et al.*, 1985). Therefore, not only is it essential to identify the various components in nuclear extracts that are responsible for poly(A) site formation, but it is also crucial to define which of these factors may be the key rate-limiting components of the polyadenylation reaction.

The recent isolation of factors from nuclear extracts that can reconstitute poly(A) site processing *in vitro* (Gilmartin *et al.*, 1988; Christofori and Keller, 1988, 1989; McDevitt *et al.*, 1988; Takagaki *et al.*, 1988, 1989; Gilmartin and Nevins, 1989) affords the possibility of defining the key rate-limiting components. Indeed, recent experiments have demonstrated that a reconstituted system employing highly purified factors accurately recapitulates the requirement for *cis*-acting sequence elements in processing (Gilmartin and Nevins, 1989). Two of these factors form a complex with the pre-mRNA, dependent on these *cis*-acting sequences. Moreover, a step-wise assembly was observed whereby an initially unstable complex was formed involving the AAUAAA element followed by the generation of a stable complex, dependent on the GU-rich downstream element (Gilmartin and Nevins, 1989). Because of the importance of these factors and sequences in processing, the generation of a sequence specific, stable complex on the RNA is suggestive of an important role in processing. To address this issue directly and as an approach to an understanding of which factor might be a rate-limiting component, we have compared the processing efficiency of several RNAs with alterations in the downstream element with the stability of the complex that forms on the RNAs. In fact, we find a direct relationship between processing efficiency and complex stability. We thus conclude that the stability of the committed processing complex may play a key role in the regulation of poly(A) site processing.

Results

An in vitro reconstituted processing system reflects in vivo sequence requirements

We have previously isolated four factors from HeLa cell nuclear extracts that were capable of reconstituting authentic poly(A) site processing *in vitro* (Gilmartin and Nevins, 1989). Two of these factors were found to be capable of specifically recognizing and complexing with a pre-mRNA. The PF2 factor, which imparts AAUAAA specificity to both

cleavage and polyadenylation, interacts with the RNA dependent on the AAUAAA element. The CF1 factor, which is required for cleavage but not polyadenylation, interacts with the RNA dependent on the PF2–AAUAAA interaction as well as the downstream sequence. Furthermore, the ternary complex (RNA–PF2–CF1) appeared to be more stable than the initial PF2–RNA complex. This finding, together with the fact that CF1 recognizes the downstream sequence element, and the likelihood that the downstream sequence may be responsible for dictating differences in poly(A) site use, based on *in vivo* transfection studies, has led us to address the possibility that the critical rate-limiting event in polyadenylation might be the CF1 interaction. Specifically, might the nature of the CF1 recognition be the determining event in the efficiency of poly(A) site use?

We have taken advantage of previous work that analyzed the requirement for the poly(A) site downstream sequence element to address the role of ternary complex formation in the efficiency of poly(A) site utilization (McDevitt *et al.*, 1986) using a pre-mRNA containing a simple downstream element from the early SV40 poly(A) site, previously shown to be essential for poly(A) site cleavage both *in vivo* (McDevitt *et al.*, 1986) and *in vitro* (Gilmartin *et al.*, 1988). The advantage in this particular system lies in the fact that previous studies have defined the downstream sequence requirement quite accurately. In particular, these experiments demonstrated that single base changes in a defined downstream element altered poly(A) site function *in vivo* (Figure 1A). These findings are also evident in a reconstituted *in vitro* processing assay employing the PF2 and CF1 factors as well as two additional factors, PF1 and CF2, that are necessary to reconstitute cleavage. As shown in Figure 1B, the reconstituted system is dependent on the AAUAAA element as indicated by the lack of processing of an adenovirus L3 RNA containing a U to C mutation in the AAUAAA element. As shown previously (Gilmartin *et al.*, 1988), two cleavage sites are utilized in the L3 poly(A) site giving rise to the two products seen in this analysis. The RNA containing the wild-type SV40 downstream sequence element (SV+) is also efficiently processed whereas the SV24 mutation, which changes two residues in the downstream element (Figure 1A), drastically reduces processing *in vitro*. Thus, two RNAs that differ by only two nucleotides are processed with a markedly different efficiency in the reconstituted system, a result very similar to that found in previous *in vivo* assays (McDevitt *et al.*, 1986).

Previous studies have demonstrated that the spacing between the AAUAAA element and the downstream element is critical for efficient poly(A) site function since displacement of the downstream element by the insertion of a spacer sequence abolished poly(A) site utilization (McDevitt *et al.*, 1986; Gil and Proudfoot, 1987; Heath *et al.*, 1990). Once again, this effect is reproduced in the reconstituted *in vitro* reaction. Using the RNA with the early SV40 downstream sequence, displacement of this sequence by 36 nucleotides (nt) of spacer sequence inserted between the cleavage site and the downstream element (Figure 1A) markedly reduced cleavage activity *in vitro* (Figure 1B). Based on the results of the assays presented in Figure 1, we conclude that the reconstituted system faithfully reproduces the *in vivo* sequence specificity for poly(A) site processing.

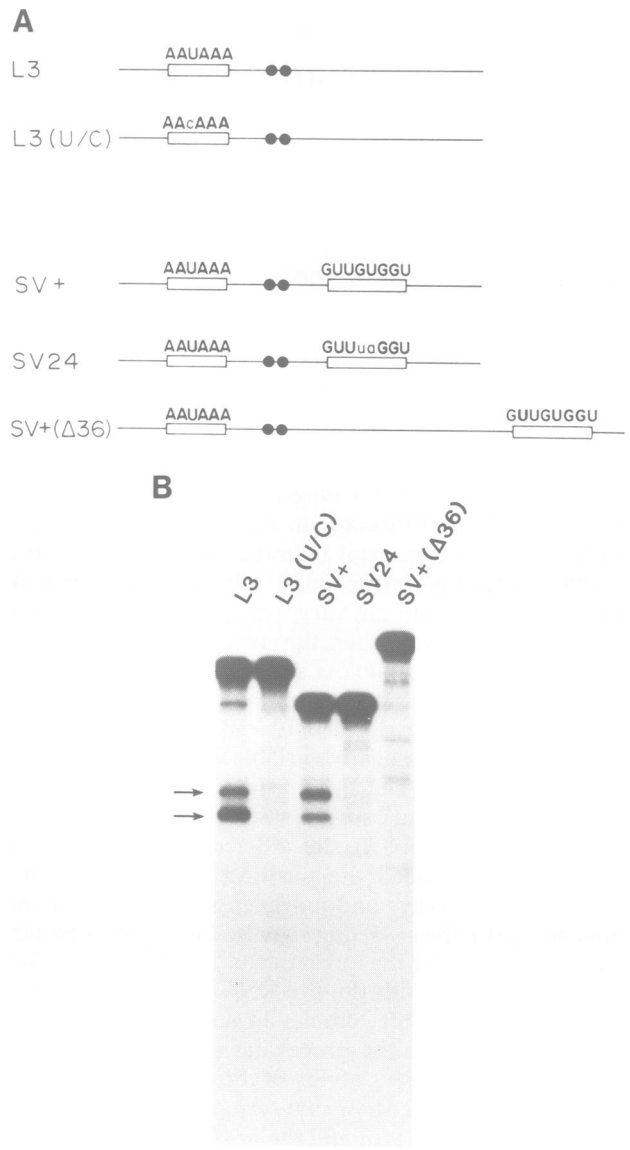

Fig. 1. *Cis*-acting sequences essential for processing in a reconstituted assay. **A.** Constructs used in the cleavage and shift assays. The sequence upstream of the cleavage site is derived from the adenovirus L3 poly(A) site and is identical in each construct with the exception of the single U to C nucleotide change in L3 (U/C). The last three constructions contain variations of the early SV40 poly(A) site as described previously (McDevitt *et al.*, 1986). SV+ contains the previously defined downstream element from the early SV40 poly(A) site; SV24 is a 2 nt change (indicated by lower case) in this element as described previously; SV+ (Δ36) contains a 36 nt insertion between the cleavage site and the downstream element. Processing takes place at two adjacent sites (solid circles), yielding products of 116 nt and 122 nt. **B.** *In vitro* 3′ cleavage assays. Each of the RNAs depicted in panel A was synthesized *in vitro* and assayed in the reconstituted system containing PF1, PF2, CF1 and CF2. The arrows indicate the two specific cleavage products.

Stable complex formation dependent on the PF2 and CF1 factors

Because the reconstituted system reflected the *in vivo* efficiencies of poly(A) site processing, we felt confident that the factors employed in these assays must interact in a way to accurately reproduce *in vivo* recognition. Using the RNA with the wild-type downstream element, we have measured

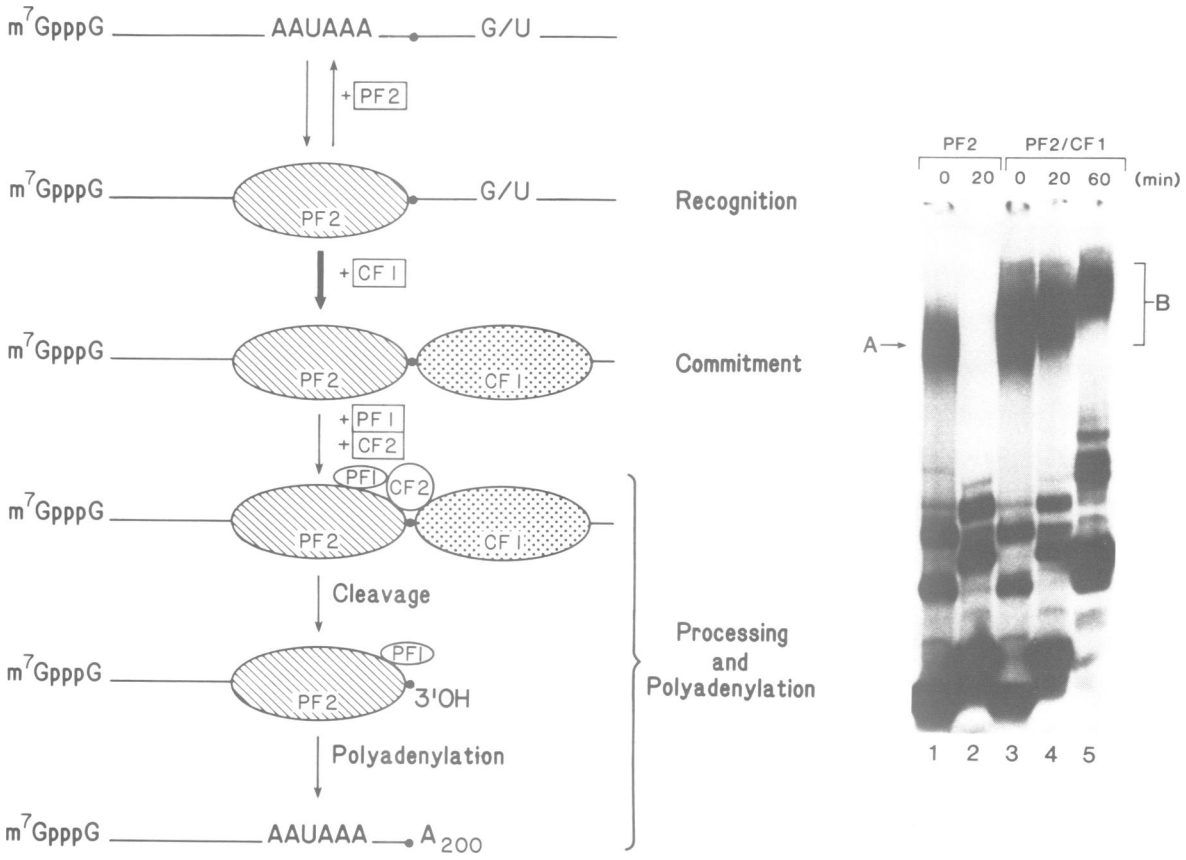

Fig. 2. Left: A summary of the previously defined RNA−protein interactions at the poly(A) site. **Right:** Dissociation rate measurements of poly(A) site complexes. RNA−protein complexes were formed with a precursor RNA containing the early SV40 downstream element (SV+ in Figure 1A) together with the PF2 factor or PF2 + CF1 factors. After allowing complex formation to reach equilibrium, samples were removed (lanes 1 and 3) and then a 100-fold molar excess of a cold RNA containing the adenovirus L3 poly(A) site was added [control experiments demonstrate that the nature of downstream sequence in the cold competitor RNA (L3 or SV40) does not alter the result]. The reactions were again sampled 20 min (lanes 2 and 4) and 60 min (lane 5) later. Samples were immediately loaded onto an acrylamide gel. Because the samples were loaded at these three time points, the complexes and the probes in the samples loaded later have not migrated as far as the early ones. The probe RNAs are the heavy bands seen at the bottom of the gel in each lane. The bands running somewhat more slowly than the probes are non-specific complexes.

the stability of the complexes formed with the previously characterized PF2 and CF1 processing factors (Gilmartin and Nevins, 1989). As shown in Figure 2A, although the PF2 factor clearly recognizes the RNA, the complex (complex A) that is formed is very unstable, having completely dissociated by 20 min. The addition of CF1, a factor essential for 3′ cleavage but not required for polyadenylation (Gilmartin and Nevins, 1989), along with PF2 generates a stable ternary complex (complex B). In contrast to the PF2 complex, the PF2−CF1−RNA complex shows no evidence of dissociation after 60 min. Moreover, further purification of CF1 activity by glycerol gradient sedimentation has shown that the activity required for reconstitution of accurate 3′ processing precisely co-sediments with the activity that combines with PF2 to yield the stable RNA−protein complex (Gilmartin and Nevins, submitted); these two activities appear to result from the same factor.

From these results as well as from previous experiments that documented the role of the factors in the processing reaction, we conclude that the interaction of the PF2 factor with the RNA, dependent on the AAUAAA sequence, represents the initial recognition event in the polyadenylation reaction (Figure 2A). This is consistent with the fact that the AAUAAA sequence appears first in the nascent transcript as well as the ubiquitous nature of the AAUAAA element; >95% of poly(A) sites contain this sequence. However, the dissociation assays indicate that this interaction is only transient and will not lead to a productive polyadenylation event, again consistent with *in vivo* data that indicate that an RNA without a downstream element is only inefficiently processed. The interaction of the CF1 factor with the RNA, dependent on the PF2−AAUAAA interaction, and dependent on the downstream sequence element, appears to commit the poly(A) site for processing. We come to this conclusion based on the stability of the complexes that form as well as the *in vivo* requirement for the downstream sequence element.

Stable complex formation reflects the efficiency of the poly(A) site

The assays of Figure 1 demonstrate that the reconstituted system reflects the *in vivo* requirements for 3′ end processing. As such, we have now used the two RNAs containing the wild-type or mutant downstream sequence to assess the relevance of stable complex formation to poly(A) site function. As shown in Figure 3A, the addition of PF2 and CF1 to each RNA resulted in the formation of the ternary

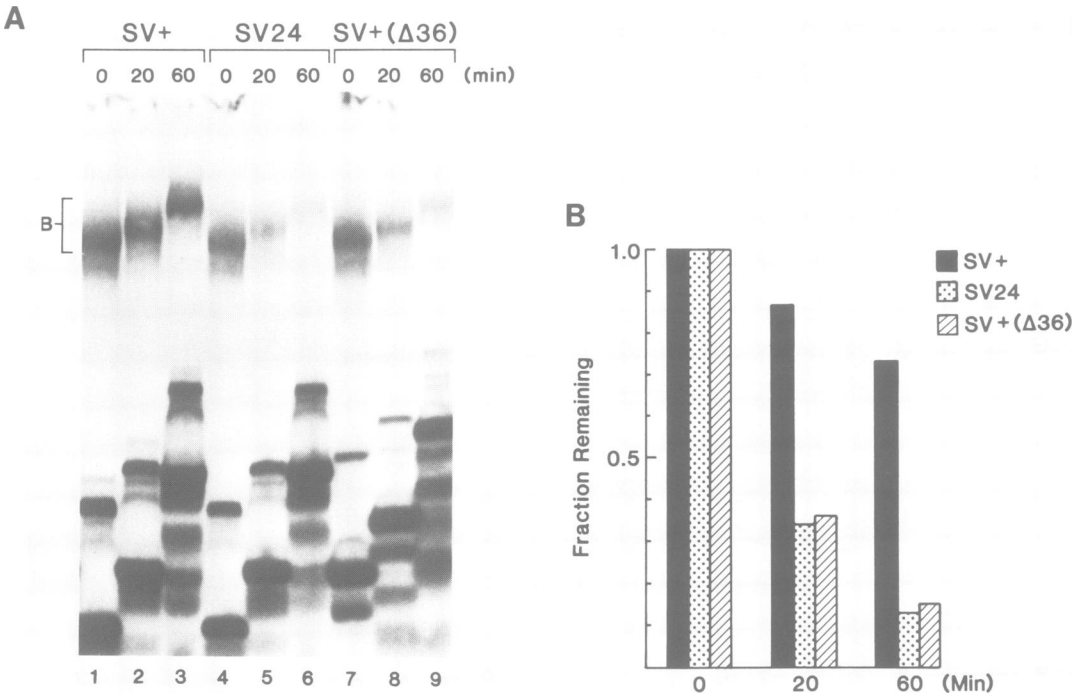

Fig. 3. The stability of the PF2−CF1−RNA complex reflects the function of the downstream element. **A**. The stability of PF2−CF1−RNA complexes formed with the SV+, SV24 and SV+(Δ36) RNAs was assessed as in Figure 2. One-third of each reaction mixture was removed after initial complex formation (lanes 1, 4 and 7). A 100-fold molar excess of a cold RNA containing the adenovirus L3 poly(A) site was then added and the reactions were sampled after further incubation times of 20 min (lanes 2, 5 and 8) and 60 min (lanes 3, 6 and 9). **B**. Relative complex stability. The region of the autoradiogram containing the B complexes shown in panel A was scanned with a densitometer and the relative values are plotted.

complex (lanes 1 and 4). However, a dissociation assay revealed a distinct difference in the stability of the complexes. Whereas the PF2−CF1−RNA complex formed on the RNA containing the wild-type SV40 downstream element was very stable, showing little evidence of dissociation after 60 min, the complex formed on the RNA with the two nucleotide change in the downstream element was much less stable. Indeed, only 13% of the complexes remained after 60 min (Figure 3B). It thus appears that a small change in RNA sequence in a downstream element, which has a significant impact on processing efficiency both *in vivo* and *in vitro*, is reflected in a reduction in the stability of the ternary complex, implying that the stability with which the purified factors interact with the downstream element determines the efficiency of processing.

That the interaction of CF1 with a downstream sequence stabilizes the PF2−RNA complex suggests the possibility of a physical interaction between the two factors. As shown in Figure 1, displacement of the GU sequence 36 nt further downstream markedly decreased processing efficiency. Analysis of the complex formed with PF2 and CF1 on this RNA indicates that the loss of cleavage activity is a reflection of a reduced stability of the ternary complex. As shown in Figure 3A, PF2 and CF1 can form a complex on this RNA (lane 7) but an off-rate measurement demonstrates that the complex is considerably less stable than the complex formed on the wild-type RNA. Thus, once again the reduction in 3′ processing efficiency, in this case as a result of an alteration in spacing of the downstream element, correlates with a reduction in the stability of the complex that forms with PF2 and CF1.

Discussion

The endonucleolytic cleavage and subsequent polyadenylation of the primary transcript that produces the mRNA 3′ terminus is a crucial event in the generation of a functional mRNA and thus an understanding of the biochemistry of this reaction is critical to a final understanding of gene regulation. Recent experiments have identified factors present in nuclear extracts that mediate the poly(A) site processing reaction (Gilmartin *et al.*, 1988; Christofori and Keller, 1988, 1989; McDevitt *et al.*, 1988; Takagaki *et al.*, 1988, 1989; Gilmartin and Nevins, 1989). The experiments reported here demonstrate that these factors can reconstitute the processing reaction with appropriate *in vivo* sequence specificity. Most importantly, the results demonstrate that the efficiency of poly(A) site use is a direct reflection of the stability of complexes formed with factors that recognize the two sequence elements that define a poly(A) site. These results strongly suggest that a crucial determinant of poly(A) site efficiency is the nature of the interaction of CF1 with the downstream element. We would suggest that an efficient poly(A) site is one that can form a stable ternary complex whereas an inefficient site forms a relatively unstable complex.

We believe it is important to consider the role of stable complex formation at polyadenylation sites in the light of experiments that demonstrate a time constraint on the polyadenylation event. First, poly(A) site recognition and processing very likely occurs on a nascent RNA transcript, shortly after the sequence has been transcribed (Nevins and Darnell, 1978). Therefore, in transcription units where

multiple poly(A) sites are utilized, the promoter-proximal site would be at an advantage over a more distal site since it would be transcribed first; a very efficient site would preclude use of a downstream poly(A) site. If the proximal site were inefficient, however, transcripts might then be processed at the second site before processing could occur at the first site. Clearly, in such a scenario the important factors would be the relative strengths of the poly(A) sites and their position within the transcription unit; the distance between the two sites should determine the time in which the first site was available for processing without competition from the second. In fact, several experiments have shown this to be the case since decreasing the distance between two tandem poly(A) sites reduces the frequency of use of the proximal site (Peterson and Perry, 1986; Galli *et al.*, 1987; Tsurushita and Korn, 1987; Denome and Cole, 1988). Taken in this context, we suggest that the relative efficiencies of the poly(A) sites are dictated by the stability of the CF1–PF2 complex that forms at the poly(A) site, as defined by the nature of the downstream GU-rich sequence element. The more stable the complex then the greater is the probability that the RNA will be processed at that location. Moreover, we also imagine that even in the absence of competing poly(A) sites, as is the predominant situation, there still exists a time constraint on the formation of the complex as the result of a competition between formation of the functional processing complex against the interaction of hnRNP proteins that block further interactions.

Based on these observations, we suggest that the regulation of poly(A) site use could depend on the differential stability of the interaction of CF1 with the processing complex. For instance, under limiting CF1 concentrations, a poly(A) site that is only poorly recognized by CF1 will not be used at high efficiency due to rapid dissociation of the PF2–CF1–RNA complex. If, however, the concentration of CF1 were to increase, then even an unstable interaction would be maintained by the concentration of the factor. Although such a mechanism does not allow a fine discrimination to be made between various poly(A) sites, and certainly does not preclude the presence of site specific factors that might be developmentally or cell type restricted, it does provide a means for modulating the use of an inefficient site. This may indeed be an underlying mechanism responsible for the switch in Ig μ mRNA production during B cell differentiation where the inefficient μ_s poly(A) site in a mature resting B cell becomes highly efficient in a secreting plasma cell (Galli *et al.*, 1987; Peterson and Perry, 1989; Weiss *et al.*, 1989). A change in CF1 from a limiting concentration to a non-limiting concentration, or an alteration of CF1 that allows a more stable ternary complex to form on the μ_s site, could account for the control.

Methods

RNA–protein complex formation

SP6 RNA synthesis was performed as described previously (Gilmartin *et al.*, 1988). PF2 was a Mono Q fraction and CF1 was a Mono S fraction, both prepared as described previously (Gilmartin and Nevins, 1989). Conditions for the assay of specific complex formation were as described (Gilmartin and Nevins, 1989). To perform the dissociation assays, an aliquot of the reaction mixture was removed at each time point, treated as described (Gilmartin and Nevins, 1989) and then loaded directly onto a 3% polyacrylamide gel.

In vitro processing

The cleavage reactions were performed as described previously (Gilmartin and Nevins, 1989) using extensively purified fractions. PF2 and CF1 were the same fractions as those used in the gel shift assays. CF2, which also contains the PF1 factor, was a Mono S fraction (Gilmartin and Nevins, 1989). Incubations were carried out in the absence of Mg^{2+} and in the presence of 20 μg/ml *Escherichia coli* tRNA and 1 mM ATP at 30°C for 90 min. RNAs were purified and analyzed in a 10% polyacrylamide gel.

Acknowledgements

We thank Lynne Dengler for her excellent assistance in the preparation of nuclear extracts. EAW was supported by a postdoctoral fellowship from the American Cancer Society (PF-3186) and GMG was supported by the Howard Hughes Medical Institute. The work was funded by a grant from the NIH (GM35984).

References

Birnstiel,M.L., Busslinger,M. and Strub,K. (1985) *Cell*, **41**, 349–359.
Christofori,G. and Keller,W. (1988) *Cell*, **54**, 875–889.
Christofori,G. and Keller,W. (1989) *Mol. Cell. Biol.*, **9**, 193–203.
Denome,R.M. and Cole,C.D. (1988) *Mol. Cell. Biol.*, **8**, 4829–4839.
Fitzgerald,M. and Shenk,T. (1981) *Cell*, **24**, 251–260.
Galli,G., Guise,J.W., McDevitt,M.A., Tucker,P.W. and Nevins,J.R. (1987) *Genes Dev.*, **1**, 471–481.
Gil,A. and Proudfoot,N.J. (1984) *Nature*, **312**, 473–474.
Gil,A. and Proudfoot,N.J. (1987) *Cell*, **49**, 399–406.
Gilmartin,G.M. and Nevins,J.R. (1989) *Genes Dev.*, **3**, 2180–2189.
Gilmartin,G.M., McDevitt,M.A. and Nevins,J.R. (1988) *Genes Dev.*, **2**, 578–587.
Heath,C.V., Denome,R.M. and Cole,C.N. (1990) *J. Biol. Chem.*, **265**, 9098–9104.
McDevitt,M.A., Imperiale,M.J., Ali,H. and Nevins,J.R. (1984) *Cell*, **37**, 993–999.
McDevitt,M.A., Hart,R.P., Wong,W.W. and Nevins,J.R. (1986) *EMBO J.*, **5**, 2907–2913.
McDevitt,M.A., Gilmartin,G.M., Reeves,W.H. and Nevins,J.R. (1988) *Genes Dev.*, **2**, 588–597.
Mclauchlan,J., Gaffney,D., Whitton,J.L. and Clements,J.B. (1985) *Nucleic Acids Res.*, **13**, 1347–1368.
Nevins,J.R. (1983) *Annu. Rev. Biochem.*, **52**, 441–466.
Nevins,J.R. and Darnell,J.E. (1978) *Cell*, **15**, 1477–1493.
Peterson,M.L. and Perry,R.P. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8883–8887.
Peterson,M.L. and Perry,R.P. (1989) *Mol. Cell. Biol.*, **9**, 726–738.
Proudfoot,N.J. and Brownlee,G.G. (1976) *Nature*, **263**, 211–214.
Sadofsky,M. and Alwine,J.C. (1984) *Mol. Cell. Biol.*, **4**, 1460–1468.
Simonsen,C.C. and Levinson,A.D. (1983) *Mol. Cell. Biol.*, **3**, 2250–2258.
Takagaki,Y., Ryner,L.C. and Manley,J.L. (1988) *Cell*, **52**, 731–742.
Takagaki,Y., Ryner,L.C. and Manley,J.L. (1989) *Genes Dev.*, **3**, 1711–1724.
Tsurushita,N. and Korn,L.J. (1987) *Mol. Cell. Biol.*, **7**, 2602–2605.
Weiss,E.A., Michael,A. and Yuan,D. (1989) *J. Immunol.*, **143**, 1046–1052.
Woychik,R.P., Lyons,R.H., Post,L. and Rottman,F.M. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3944–3948.

Received on August 30, 1990; revised on October 22, 1990