Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 5854–5857, August 1986
Biochemistry

Autogenous regulation of splicing of the transcript of a yeast ribosomal protein gene

(intron/cDNA clone/gene fusion/*Saccharomyces*)

MARIANA D. DABEVA*, MARTHA A. POST-BEITTENMILLER, AND JONATHAN R. WARNER†

Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY 10461

Communicated by Harry Eagle, April 25, 1986

ABSTRACT **The gene for a yeast ribosomal protein, *RPL32*, contains a single intron. The product of this gene appears to participate in feedback control of the splicing of the intron from the transcript. This autogenous regulation of splicing provides a striking analogy to the autogenous regulation of translation of ribosomal proteins in *Escherichia coli*.**

In the yeast *Saccharomyces cerevisiae*, very few genes have introns. Yet 26 of 30 cloned ribosomal protein genes contain a single intron near the site of initiation of translation (reviewed in ref. 1). The ribosomal proteins represent a unique group whose synthesis is regulated coordinately to yield nearly equimolar amounts of each (reviewed in ref. 2). These two features suggest that the introns in the ribosomal protein genes may play a role in the coordinate regulation of their synthesis. Support for such an idea comes from the observation that when either of two ribosomal protein genes, *RPL32* or *CYH2* (coding for rpL29), are present in multiple copies, the cell accumulates a large excess of the unspliced transcript of that gene (3).

Based on these considerations, we suggest a hypothesis that postulates that the product of a ribosomal protein gene can regulate the splicing of its own transcript (Fig. 1). Thus, for newly synthesized ribosomal proteins there is a competition between newly transcribed ribosomal RNA and unspliced mRNA. An excess of ribosomal protein over ribosomal precursor RNA could lead to its binding to the unspliced transcript and preventing the formation of mature mRNA. This is essentially a eukaryotic analogue to the autogenous regulation of translation of ribosomal proteins, which has been demonstrated to occur in *Escherichia coli* (4).

Bozzoni *et al.* also found that when *Xenopus* oocytes were injected with a very large number of copies of the gene for *Xenopus* ribosomal protein L1, there was an accumulation of transcripts whose second intron was aberrantly spliced and whose third intron was not spliced at all. No such effect was found when the gene for *Xenopus* ribosomal protein L14 was injected. They suggest that L1 can regulate the splicing of its transcript (5).

It is important to distinguish specific regulation of splicing on the one hand from inefficient or aberrant splicing on the other. To make such a distinction, we propose and test a number of predictions of the hypothesis described in Fig. 1. The data show that the gene *RPL32* fulfills each of these predictions, leading us to conclude that ribosomal protein L32 can inhibit, directly or indirectly, the splicing of the transcript of its gene.

MATERIALS AND METHODS

Methods. *S. cerevisiae*, strain S150-2B (*MATa, trp 1-289, his 3-1, ura 3-52, leu 2-3,112*), was transformed with 2μ-based vectors YEp24′ (3) or pMAC 561 (6) containing the indicated constructs. Transformations were carried out, cellular DNA and RNA were prepared, and the copy number of plasmids was determined by quantitative Southern analysis as described (3).

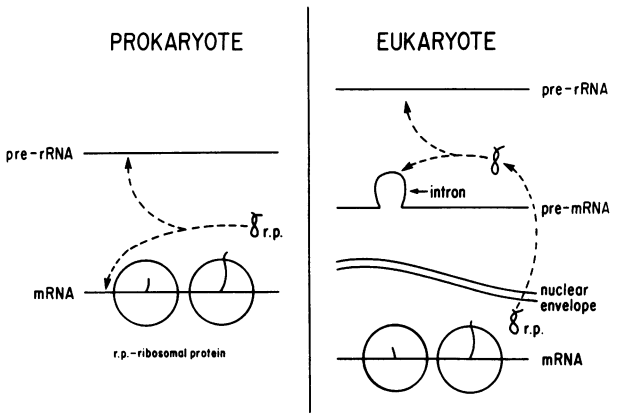
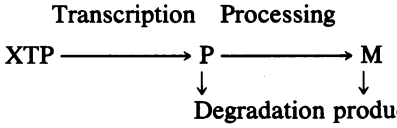


FIG. 1. Model of the autogenous regulation of splicing. See text for details.

Primer extension analysis was carried out essentially as described by Pikielny and Rosbash (7) using primers defined in Fig. 2. Twenty nanograms of oligodeoxynucleotide, 5′ end-labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase, was added to 25 μg of total yeast RNA in 0.1 M Tris·HCl, pH 8.3/0.14 M KCl, in a final vol of 8 μl. The samples were heated for 3 min at 90°C and transferred to a 41°C bath. After 5 min, 12 μl of the above buffer containing 15 mM $MgCl_2$, 3 mM dithiothreitol, 0.45 mM each dNTP, and 12 units of reverse transcriptase was added. The extension reaction was carried out for 2 hr at 41°C and was stopped by ethanol precipitation. The nucleic acids were dissolved in 5 μl of RNase (1 mg/ml) and 1.25 μl of the sample was mixed with 3 μl of loading buffer, treated for 3 min at 95°C, and loaded on a 7% polyacrylamide/7 M urea sequencing gel. After electrophoresis, the gels were dried and exposed at −70°C to Kodak X-Omat film.

RESULTS

Accumulation of Unspliced Transcripts. One can analyze the flow of nucleotides through a given mRNA with the following diagram:



The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*On leave from the Institute of Cell Biology, Bulgarian Academy of Sciences, Sofia, Bulgaria.
†To whom reprint requests should be addressed.

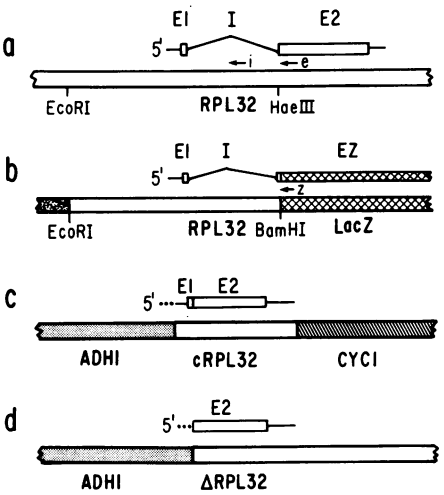


FIG. 2. Constructs used in the analysis of splicing of the transcript of *RPL32*. (*a*) The gene for *RPL32* is presented. The transcript, shown above, consists of three parts: a 58-nucleotide leader and the first exon consisting only of AUG, an intron of 230 nucleotides, and a second exon of 414 nucleotides. [Details of the structure and sequence of *RPL32* will be presented elsewhere (M.D.D. and J.R.W., unpublished data).] Fragments including the whole gene and either 600 or 900 nucleotides upstream of the transcription initiation site were cloned into the vector YEp24′ (12), selectable with the *URA3* gene, and introduced into yeast cells. The location of the intron (i) and exon (e) primers used to analyze transcripts are shown. (*b*) The *RPL32/lacZ* fusion gene was constructed by cutting at an *Hae* II site in the first codon of the second exon, introducing a *Bam*HI linker, and inserting an *Eco*RI/*Bam*HI fragment in front of the *lacZ* gene in pSEY101 (10), selectable with the *URA3* gene. The product of the fusion gene is predicted to contain only the NH_2-terminal methionine from the *RPL32* gene. The primer used to analyze the transcripts of the fusion gene (z) consists of sequences on both sides of and including the *Bam*HI linker. (*c*) Two cDNA clones of *RPL32*, in the 2μ-based vector pMAC561 selectable with the gene *TRP1*, were isolated from a library kindly provided by G. McKnight (6). They differed slightly at the 5′ and 3′ ends of the insert but, by sequence analysis, both contained all the coding sequences of *RPL32* as well as the splice junction predicted from the sequence of the gene (M.D.D. and J.R.W., unpublished data). The *ADH1* promoter and origin of transcription drive the transcription of *cRPL32* and the *CYC1* termination region provides an auxiliary termination signal (6). (*d*) An inactive derivative of the *RPL32* gene: the promoter and 5′ end of the transcript are derived from the *ADH1* gene, which is fused to the *RPL32* gene just downstream of the intron. It is thus missing not only the intron but also the initiation codon, AUG, provided by the first exon. It is thus unable to code for the synthesis of ribosomal protein L32. We have confirmed by *in vitro* translation that RNA from cells carrying ΔRPL32 codes for excess L32 while RNA from cells carrying cRPL32 does not (data not shown).

where P represents the concentration of unspliced transcripts and M is the concentration of mature (i.e., spliced) transcripts. In the usual situation, a yeast cell contains very little precursor, so that the ratio $P/M << 1$ (8, 9). In the presence of multiple copies of the genes *RPL32* or *CYH2*, $P/M \geq 1$ for the transcript of the genes present in excess, although for transcripts derived from other genes this ratio remains small (3). Therefore, the increase in unspliced transcript is specific for the multicopy gene and is not due to a general overloading of the splicing machinery of the cell.

The hypothesis we wish to test states essentially that the rate of processing of a transcript can be modulated by the product of its gene. One prediction is that increasing the transcription of *RPL32*, which is proportional to the copy number of the gene, will lead to an accumulation of precursor proportionately greater than the increase in copy number. This is true because the product of a single copy of the gene is consumed in the synthesis of ribosomes, while the product of any additional copies will be in excess and thus available to inhibit splicing. This prediction was tested by comparing, using primer extension, the RNA of cells carrying a single copy of *RPL32* (Fig. 3, lanes a and c) with that from cells carrying many copies (lanes b and d). It is clear that in cells carrying multiple copies of *RPL32* the level of unspliced transcripts is increased >100-fold, while the copy number of the plasmid carrying *RPL32*, estimated by Southern analysis of DNA isolated from the cells, is 12–15 (data not shown). Because the primer extension proceeds to the 5′ end of the molecule, the accumulated transcripts are not in the lariat form (10). Therefore, it is the first step of the splicing reaction that is inhibited. The intron of *RPL32* contains the invariant

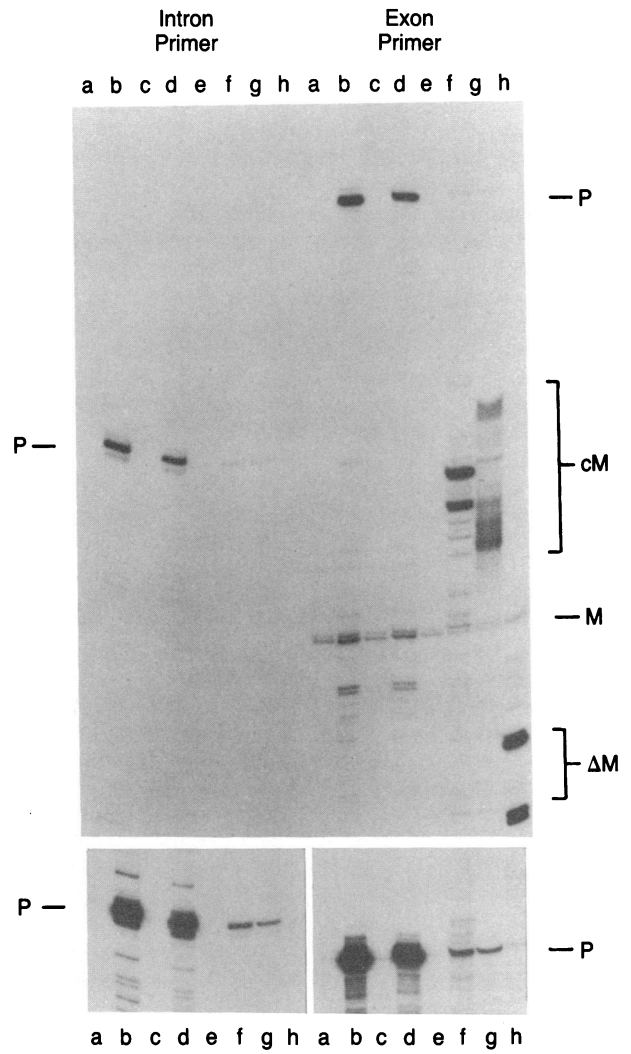


FIG. 3. Primer-extension analysis of transcripts from the gene for *RPL32*. Primer extension was carried out as described. Two synthetic oligonucleotides were used as primer (see Fig. 2): i, a 16-mer with a sequence complementary to nucleotides 136–151 of the intron; and e, an 18-mer with a sequence complementary to nucleotides 29–46 of exon 2. Oligonucleotides were synthesized by solid-phase DNA synthesis on an Applied Biosystems Model 380-A. The RNA was derived from strain S150-2B carrying the following plasmids: lanes a, YEp24′, vector for the intact genes; lanes b, pYERPL32, *RPL32* (to nucleotide −900 from the AUG) in YEp24′ (12 copies per cell); lanes c, no plasmid; lanes d, pYERPL32-A, *RPL32* (to nucleotide −600) in YEp24′ (15 copies per cell); lanes e, pMAC561, vector for cDNA genes; lanes f, pcRPL32, *cRPL32* in pMAC561; lanes g, pc′RPL32, *cRPL32* in pMAC561 (an independent isolate); lanes h, pΔRPL32, *RPL32* with 1st exon and the intron deleted, in pMAC561. The products derived from unspliced (P), spliced (M), cDNA (cM), and Δ (ΔM) RNAs were identified by comparison with size markers run in adjacent lanes (not shown). (*Insets*) Longer exposures of the precursor region of the gels.

UACUAAC sequence, which determines the 3′ splice site (11). The 5′ splice site is unusual, however, in having the sequence GUCAGU instead of GUAUGU (M.D.D. and J.R.W., unpublished data). Although this variation may be involved in the regulation of splicing, Fig. 2 shows that it does not inhibit the rate of splicing under normal conditions.

Splicing of a Fusion Transcript. If the rate of splicing is affected only by the product of the gene, then an increase in transcription not accompanied by an increase in the gene product should not lead to an accumulation of precursor. Such a situation was generated by constructing an *RPL32/lacZ* fusion gene in which the first exon, the intron, and one nucleotide of the second exon of *RPL32* were fused to the coding sequences of the *E. coli* β-galactosidase gene (Fig. 2; ref. 13). Since the first exon codes only for the initiating methionine, no trace of L32 remains in the product of the fusion gene. When this gene is maintained in the cell on a multicopy plasmid (15–20 copies per cell), the ratio of P/M is ≈0.3 (Fig. 4, lane Z). In contrast, the ratio of P/M of the transcript of the *RPL32* gene itself on a multicopy plasmid is >2 (Fig. 2, lanes b and d).

By the same reasoning, we predict that when the fusion gene is present in the same cell with a multicopy plasmid carrying *RPL32*, the splicing of the fusion gene will be inhibited. Such is the case in cells carrying both plasmids (Fig. 4, lane Z+L32), where P/M for the fusion gene increases to 3.0. These two results, therefore, fulfill the predictions of the model.

Effect of a cDNA Clone on Splicing of the *RPL32* Transcript. If the rate of splicing is affected by the product of the gene, then an increase in the product of the gene without an increase in the production of full-length transcripts should lead, nevertheless, to an inhibition of splicing. We now show that this is the case both for the transcript of the endogenous gene and for the transcript of the *RPL32/lacZ* fusion gene. Two cDNA clones for *RPL32* were isolated from a library kindly provided by G. McKnight (6) (Fig. 2*c*). The two clones differ slightly at both the 5′ and 3′ ends, but sequence analysis showed that each contains the complete sequence of the *RPL32* mRNA (M.D.D. and J.R.W., unpublished data). Transcription is driven by the *ADH1* promoter (14).

When either of these two cDNA clones is introduced into a cell, it gives rise to a large amount of transcripts, with somewhat variable 5′ ends (Fig. 3, lanes f and g). The genomic *RPL32* is the only gene with the information to produce a transcript containing the intron. Its transcripts accumulate as unspliced precursor (lanes f and g). A comparison with cells carrying the vector for the cDNA clones (lanes e) suggests that the level of precursor is increased 10- to 25-fold in each case.

Similarly, the presence of the cDNA clones leads to the accumulation of unspliced transcripts derived from the *RPL32/lacZ* fusion gene (Fig. 4, lane Z+cL32). Comparison with control cells (lane Z) or cells carrying the cDNA vector (not shown) suggests that the P/M ratio is increased by ≈10-fold. A similar value was obtained from cells carrying the other cDNA clone. These results again confirm the predictions of the hypothesis.

Effect of a Defective *RPL32*. The results presented above demonstrate that the cDNA clone affects splicing but do not directly establish the effector, which could be the cDNA itself, its transcript, or its product. To establish more clearly that the gene product is the effector, we constructed a defective *RPL32*, which lacks the initiator AUG and the intron (Fig. 2). Thus, it is unable to code for the L32 protein. When this defective gene construct is introduced into a cell, it is actively transcribed (Fig. 3, lane h: ΔM), but it leads to no accumulation of the unspliced precursor either of the genomic *RPL32* (lane h) or of the *RPL32/lacZ* fusion gene (Fig. 4, lane Z+ΔL32). Therefore, we conclude that the accumulation of unspliced precursor, which occurs in the presence of excess copies of either normal or cDNA *RPL32* genes, is due to the gene product itself.

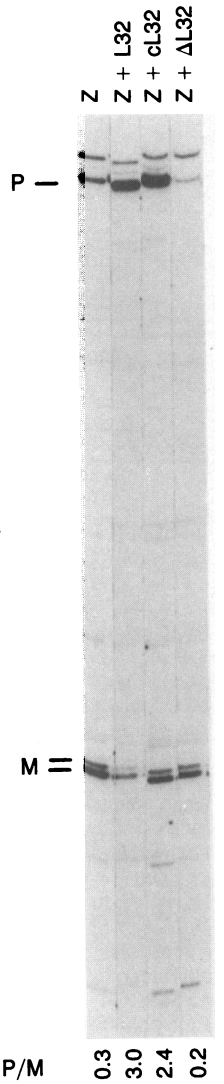


FIG. 4. Primer extension analysis of transcripts from the *RPL32/lacZ* fusion gene. Primer extension was carried out as described with a primer that spanned the junction between *RPL32* and *lacZ*, to score only fusion transcripts (Fig. 2). The RNA was derived from strain S150-2B carrying the following plasmids. In each case the copy number of the fusion plasmid was >20, and the ratio of the fusion plasmid to the other was constant. Lanes: Z, RPL32/Z in pSEY101, maintained under uracil selection; Z+L32, RPL32/Z in pSEY101 and RPL32 in YEp24′, maintained under uracil selection; Z+cL32, RPL32/Z in pSEY101 and cRPL32 in pMAC561, maintained under uracil and tryptophan selection; Z+ΔL32, RPL32/Z in pSEY101 and ΔRPL32 in pMAC561, maintained under uracil and tryptophan selection. The products derived from unspliced (P) and spliced (M) RNAs were identified by comparison with size markers run in adjacent lanes. The ratio of P/M as determined by quantitative densitometry of several exposures is indicated below the lanes.

DISCUSSION

The data presented in Figs. 3 and 4, together with data presented previously (3), demonstrate that the accumulation of unspliced transcript of L32 is not due to a general overloading of the splicing apparatus and is not due to a specific overloading of the splicing reaction for the transcript of *RPL32*. The accumulation depends on the overproduction of an mRNA coding for an intact L32, and thus presumably on overproduction of L32 itself. The results are consistent

with the predictions derived from the model presented in Fig. 1. We conclude that ribosomal protein L32 inhibits the splicing of the transcript of its gene.

Since cells carrying a multicopy plasmid with the gene *CYH2* accumulated large amounts of unspliced precursor, we carried out similar experiments to ask if that gene fit the model. However, the increase in amount of precursor, while substantial, was no greater than the increase in copy number (figure 2*B* of ref. 3). Furthermore, the presence of a functional cDNA clone of *CYH2* did not lead to a detectable increase in unspliced transcripts of the endogenous gene (data not shown). We conclude that the splicing of the transcript of *CYH2* is relatively slow but is not affected by the presence of its product, ribosomal protein L29. These results demonstrate that such experiments can distinguish regulation of splicing from inefficiency of splicing. One non-ribosomal gene of *Saccharomyces* with introns is the *MATα1* gene, which codes for a protein regulating mating type. One of the two introns is removed from its transcript slowly (15). It will be interesting to determine whether this is due to the inefficiency or the regulation of its splicing.

The splicing reaction is so complex (16, 17), and the number of RNA and protein molecules involved is so large, that it is premature to speculate about precise mechanisms by which its regulation might occur. Nevertheless, the results presented above suggest that L32 itself interacts with the transcript of its gene to inhibit the first step in its splicing—i.e., cleavage at the 5′ splice site and formation of a lariat structure (7, 18). Because of the effect of the cDNA clones on the *RPL32/lacZ* fusion, we can limit the site of interaction to the 5′ exon and the intron. The sequence of this region (M.D.D. and J.R.W., unpublished data) has been compared with the sequence of yeast 25S rRNA to search for regions of homology that might compete for the binding of L32. Some homologies of 8–10 nucleotides were found. However, since the binding site is presumably a three-dimensional RNA structure (4), one-dimensional homology may not be the appropriate comparison. We will need to define more precisely the region on the *RPL32* transcript that is involved as well as the site on 25S rRNA to which L32 binds. It remains to be seen whether those situations in which the splicing of a transcript varies in a developmental-specific or tissue-specific manner—for example, in the cases of immunoglobulins (19) or troponin T (20)—are brought about by an equivalent mechanism.

The analyses of the dynamics of mRNA transcripts presented above have ignored the turnover of precursor and mature RNAs, for the simple reason that they are difficult to evaluate experimentally. Nevertheless, they clearly play a role in determining the concentration of precursor and mature molecules. For instance, we presume that the mRNA derived from the *RPL32/lacZ* fusion gene is relatively unstable since the gene does not have a yeast poly(A) addition site. The instability of the mature mRNA may be responsible for the P/M ratio of the transcript of the fusion gene (Fig. 4, lane Z) being higher than that for the RNA from the endogenous *RPL32* gene (Fig. 3, lanes a and c).

Let us now compare the regulation of synthesis of ribosomal proteins in *E. coli* and in *S. cerevisiae* (Fig. 1). They use the same mechanism (binding a gene product to its transcript in competition with pre-rRNA) for the same end (to limit the synthesis of new protein). But the limitation is at a different level: in *E. coli*, translation; in *S. cerevisiae*, splicing. The attraction of this comparison is that it makes sense in evolutionary terms. As eukaryotes developed from prokaryotes, they lost the intimate geographical relationship between transcription and translation that is the basis for the regulation of ribosomal protein synthesis. As an alternative, without having to abandon the binding of protein to its mRNA as the determinant of specificity, they developed a relationship between binding and mRNA production. This may, therefore, be the reason that the ribosomal protein genes of yeast are conspicuous for their introns.

Yet, of several yeast ribosomal protein genes that have been reintroduced into cells on multicopy plasmids (3, 21), only *RPL32* shows clear evidence of regulation at the level of splicing. Therefore, one is left to speculate on the role of the introns in the other genes. Are they remnants of evolution or are they regulators during nonlogarithmic phases of the growth cycle, or during differentiation—e.g., sporulation or germination?

We are grateful to Dr. Gary McKnight for the cDNA library; to Dr. Scott Emr for the *lacZ* fusion vector pSEY101; and to Drs. Elaine Elion, Stewart Johnson, Julius Marmur, and Arthur Skoultchi for comments on the manuscript. This work was supported by grants from the American Cancer Society (NP-474-0) and the National Institutes of Health (GM25532 and CA13330). The oligonucleotides were from a facility funded by an equipment grant from the National Science Foundation (PCM-8400114). M.A.P.-B. is a Fellow of the American Cancer Society.

1. Leer, R. J., van-Raamsdonk-Duin, M. M., Hagendoorn, M. J., Mager, W. H. & Planta, R. (1984) *Nucleic Acids Res.* **12,** 6685–6700.
2. Fried, H. M. & Warner, J. R. (1984) in *Recombinant DNA and Cell Proliferation*, eds. Stein, G. S. & Stein, J. L. (Academic, New York), pp. 169–192.
3. Warner, J. R., Mitra, G., Schwindinger, W. F., Studeny, M. & Fried, H. M. (1985) *Mol. Cell. Biol.* **5,** 1512–1521.
4. Nomura, M., Gourse, R. & Baughman, G. (1984) *Annu. Rev. Biochem.* **53,** 75–118.
5. Bozzoni, I., Fragapane, P., Annesi, F., Pierandrei-Amaldi, P., Amaldi, F. & Beccari, E. (1984) *J. Mol. Biol.* **180,** 987–1005.
6. McKnight, G. L. & McConaughy, B. L. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 4412–4416.
7. Pikielny, C. W. & Rosbash, M. (1985) *Cell* **41,** 119–126.
8. Rosbash, M., Harris, P. K., Woolford, J. L., Jr., & Teem, J. L. (1981) *Cell* **24,** 679–686.
9. Fried, H. M., Pearson, N. J., Kim, C. H. & Warner, J. R. (1981) *J. Biol. Chem.* **256,** 10176–10183.
10. Pikielny, C. W., Teem, J. L. & Rosbash, M. (1983) *Cell* **34,** 395–403.
11. Langford, C. J. & Gallwitz, D. (1983) *Cell* **33,** 519–527.
12. Pearson, N. J., Fried, H. M. & Warner, J. R. (1982) *Cell* **29,** 347–355.
13. Douglas, M. G., Geller, B. L. & Emr, S. D. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 3983–3987.
14. Russell, D. W., Smith, M., Williamson, V. M. & Young, E. T. (1983) *J. Biol. Chem.* **258,** 2674–2682.
15. Miller, A. M. (1984) *EMBO J.* **3,** 1061–1065.
16. Brody, E. & Abelson, J. (1985) *Science* **228,** 963–967.
17. Grabowski, P. J., Seiler, S. R. & Sharp, P. A. (1985) *Cell* **42,** 345–353.
18. Domdey, H., Apostol, B., Lin, R. J., Newman, A., Brody, E. & Abelson, J. (1984) *Cell* **39,** 611–621.
19. Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. & Hood, L. (1980) *Cell* **20,** 313–319.
20. Medford, R. M., Nguyen, H. T., Destree, A. T., Summers, E. & Nadal-Ginard, B. (1984) *Cell* **38,** 409–421.
21. Abovich, N., Gritz, L., Tung, L. & Rosbash, M. (1985) *Mol. Cell. Biol.* **5,** 3429–3435.