Hypothesis

ATTENUATION AS A GENERAL MECHANISM FOR THE REGULATION OF DIFFERENTIAL GENE TRANSCRIPTION IN EUKARYOTES

J. A. REMINGTON

Département de biologie, Université Laval, Québec GPK 7P4, PQ Canada

Received 4 January 1979
Revised version received 13 February 1979

1. Introduction

As an alternative to hypotheses of eukaryotic gene regulation which involve the association of DNA and regulatory molecules, I wish to propose a model involving the specific binding of regulatory molecules to messenger RNA precursors (pre-mRNA) to control rates of differential gene transcription. This model, based on that of Lee and Yanofsky [1] for the *trp* operon of *Escherichia coli*, emphasizes the importance of the noncoding 'leader' sequences at the 5′-termini of eukaryotic pre-mRNAs in the 'attenuation' of transcription. According to the model, the secondary and/or tertiary structure at the 5′-end of the nascent pre-mRNA determines the degree of early transcription termination. Formation of a specific termination complex will result in the abortive arrest of transcription and release of uncompleted, nascent transcripts from the template. A modulation of the secondary or tertiary structure at the 5′-end of the pre-mRNA by specific regulatory molecules (antiattenuators) which bind to the pre-mRNA would alter the degree of early transcription termination and result in antiattenuation. Antiattenuation would consequently involve a change in the elongation efficiency of transcription and in the overall rate of transcription (fig.1). This model, and its elaborations, can be used to explain a number of disparate experiments on gene transcription and regulation in eukaryotes.

It is now evident that in eukaryotic cells differential gene transcription is one major regulatory step in a hierarchy of controls which also includes processing of mRNA precursors and selective transport of mRNA sequences to the cytoplasm [2,3]. Recent studies show that processing of the pre-mRNA involves excision of noncoding intervening sequences and splicing together of the coding sequences [4,5]. The pre-mRNA

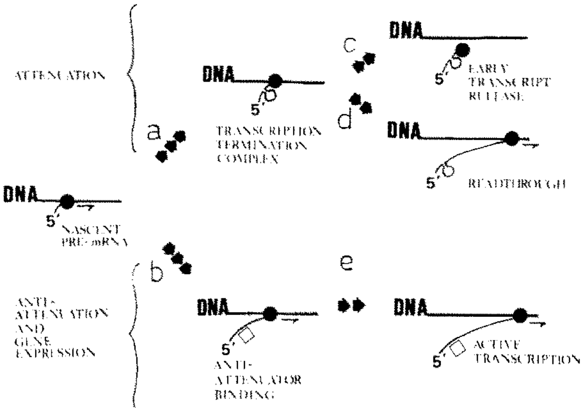

Fig.1. Schema of attenuation and antiattenuation in eukaryotic nuclei according to the model described in the text. RNA polymerase (•); antiattenuator molecule (□). Changes in the conformation of the RNA are indicated by the loop. Attenuation (a): nonspecific initiation of transcription can result in early transcription termination and the release of abortive transcripts (a,c). A certain percentage of transcription readthrough (d) will nevertheless occur whereby complete, though not necessarily functional, transcripts are formed. This readthrough may result in a nonfunctional, 'non-pre-mRNA' population described by Wold et al. [47]. Antiattenuation (b): Initiation of transcription is followed by the binding of regulatory molecules, termed antiattenuators, to the nascent pre-mRNA. The resulting modulation of the conformation of the transcript near its 5′-terminus will prohibit the formation of the transcription–termination complex. Active transcription (e) will ensue.

also contains a 5′ noncoding leader sequence and a 3′ noncoding sequence, both of which are retained in the functional mRNA molecule [6,7].

In the past decade supporting evidence for specific differential transcription as the central mechanism of gene expression in eukaryotes has come primarily from chromatin transcription studies in vitro and from the measurement of RNA sequence complexity in vivo. A number of regulatory schemes [8–10] modeled to varying extents after the *lac* operon of *E. coli* have been proposed in which the initiation of gene transcription is specified by regulatory molecules. Some of these models have been shown to be genetically unreasonable [11]. It has been claimed that the in vitro transcription of tissue-specific RNA from chromatin reconstituted in the presence of tissue-specific, nonhistone nuclear proteins (NHCPs) provides evidence for the central role of specific initiation of gene transcription and for the importance of NHCPs in the regulation of eukaryotic gene expression [12–14]. It is now known, however, that these studies have been complicated generally by the fact that exogenous *E. coli* RNA polymerase was used; this enzyme transcribes from tissue-specific RNA contaminating the chromatin preparations [15]. The more recent experiments cast serious doubt upon the validity of interpretation of the earlier in vitro transcription experiments (reviewed in [16]). Zazloff and Felsenfeld [15] have shown that there is, in fact, little or no differential transcription of mRNA globin sequences when reticulocyte chromatin is transcribed in vitro using mercurated nucleotides. On the other hand, Towle et al. [17], using the oviduct system with mercurated nucleotides, found a stimulation of transcription in vitro of the ovalbumen sequence after estrogen administration. But the preferential transcription was only 10-times more than oviduct globin gene transcription and only 4-times more than ovalbumen gene transcription from reticulocyte chromatin. Other studies of transcription in vitro and in vivo show a stimulation of tissue-specific transcription after hormone treatment of only about 1 order of magnitude [18–20].

The more recent measurements of pre-mRNA sequence complexity in vivo by RNA–DNA hybridization studies are also consistent with the in vitro data indicating relatively small transcriptional changes during induction of gene expression. Nuclear RNAs transcribed from the (gene-containing) single-copy DNA of different embryonic stages of sea urchins have practically the same sequence complexities at times when cytoplasmic sequence complexities are very different [21]. Transcriptional differences between different stages of the mammalian cell cycle may also be less important [22] than supposed [23]. Measurements of sequence complexity of the nuclear RNA (pre-mRNA) transcribed from single-copy DNA of different tissues are revealing an interesting pattern of transcriptional differences. Nuclear RNA sequences are extensively shared between different tissues [24–28]. Chikaraishi et al. [27] have found that mixtures of nuclear RNAs from different tissues, which are simultaneously annealed to single-copy DNA, hybridize to the same extent as the more complex of the RNA preparations isolated from one of the tissues. For instance, all kidney sequences (i.e., > 95%) are found among the more complex liver sequences, and all liver sequences (i.e., > 94%) are found among the yet more complex population of brain sequences. The least abundant RNAs possess the greatest sequence complexity. In fact, the average frequency of occurrence of each nonrepetitive nuclear RNA is under one (0.07–1.0) copy per cell [21,27,28]. The frequency distribution of heterogeneous nuclear (hn) RNA abundance classes is not the same for different tissues [25–27]. Also, transcriptional stimulation or inhibition in a single cell type may change the frequency distributions of classes of pre-mRNA molecules [29]. However, no large changes over 2 or 3 orders of magnitude, such as seen with some bacterial operons, and no pattern of complete disappearance of primary transcripts have been demonstrated in eukaryotic somatic cells. The above data are part of an accumulating body of evidence which indicates that transcriptional differences between cells of different tissues, developmental or cell-cycle stages, or hormonal states may be quantitative more than qualitative. It is quite possible that no specific or stringent ‘turning on’ (initiation) or ‘turning off’ (repression of initiation) of specific genes occurs at the level of transcription in somatic cells of eukaryotes.

2. Attenuation in bacteria and eukaryotes

The attenuation model of gene regulation proposed here is cotranscriptional in that a modulation of secondary and/or tertiary structure of the 5′-terminus of the pre-mRNA by antiattenuator molecules which bind to the pre-mRNA accounts for changes in the rate of transcription of selected genes. Attenuator sequences and leader RNAs involved in early transcription termination have been demonstrated in several bacterial operons [30–33]. Lee and Yanofsky [1] have proposed a model in which the stability of the secondary structure at the 3′-end of the leader sequence of mRNA influences transcription. In its simplest form their model proposes that the rate of ribosome movement along the RNA will shift the RNA intramolecular base configurations and effect the 'stem and loop' secondary structures, either allowing (rapid ribosome movement) or prohibiting (ribosome arrest or slow ribosome movement) the formation of termination complexes. There is a certain frequency of readthrough of transcription beyond the termination site under conditions which normally would be expected to result in transcription termination. (Transcription readthrough is also a feature of the eukaryotic model introduced here.) The readthrough frequency can be quite variable (from 5–76%) depending upon the operon, bacterial species, and presence of mutations in the attenuator sequence [1,30,31]. Single base-pair alterations in the *E. coli trp* operon leader region, which destabilize 'stem and loop' structures, relieve transcription termination at the *trp* attenuator [31]. The rate of ribosome movement and ribosome arrest are determined, in turn, by levels of certain charged tRNAs (for instance, charged tRNA^{Trp} for the *trp* operon). Some mutations that decrease the levels of specific, charged tRNAs by 15–60% have a 3–10-fold effect on gene expression apparently by affecting attenuation [32].

There are several differences between the bacterial model and the eukaryotic model proposed here (fig.1). There is no coupling of transcription and translation in eukaryotic cells as in bacterial cells. Consequently, ribosome movement and arrest cannot modulate attenuation. Instead, antiattenuation must be directed by the binding to pre-mRNA of other regulatory molecules, termed antiattenuators, which are either protein, RNA or some combination of the two species of molecules. It is likely that the antiattenuator has a protein component. It is known that NHCPs and hormone receptors, which are proteins, can have a direct effect on transcription [14,34]. There is, additionally, a body of evidence which suggests that small, nuclear RNA species may be implicated directly in gene transcription [35–41]. The binding of the antiattenuator need not necessarily occur at the 5′-terminal leader sequence of the pre-mRNA. However, the binding is pictured as occurring primarily at the 5′-terminus for the sake of clarity and economy (fig.1).

An ancillary, post-transcriptional feature of this model is the proposition that the binding of the pre-mRNA to the small ribosomal subunit would require an appropriate, accessible conformation at the 5′ ribosome binding site of the pre-mRNA. A modulation of the 5′ leader conformation would theoretically change the ability of the small ribosomal subunit to bind to the pre-mRNA. According to the model, therefore, another function of antiattenuator molecules, in modulating secondary and tertiary structure, may be the induction of a specific conformation enabling attachment to the small ribosomal subunit. Messenger sequences would thus be selected from the mass of hnRNA for ribosome binding. Hydrogen bonding of a sequence near the 5′-terminus of messenger RNA to a complementary sequence near the 3′-terminus of the 18 S ribosomal RNA has been suggested as a mechanism of ribosome attachment [42]. An impressive degree of such complementarity has been shown to exist for some, but not all, eukaryotic messages [43]. The proposition here is that it is not the primary sequence which is directly recognized for attachment but a specific secondary or tertiary conformation induced by the binding of an antiattenuator. Thus, a common mechanism may account for relief of abortive transcription termination and for the selection of messenger sequences for ribosome attachment. Antiattenuation of transcription and ribosome binding may be coupled to modulation of the same secondary and tertiary structure on the pre-mRNA. On the other hand, very different binding sites for the antiattenuation of transcription and the facilitation of ribosome attachment could exist, in which case the two phenomena would not be coupled.

3. Discussion

Attenuation and antiattenuation in eukaryotes is presented here as a major regulatory step in a hierarchy of controls by which genes are differentially expressed. To what extent does the attenuation model of transcriptional regulation account better for the eukaryotic transcription data than existing operon-like, initiation models? Several points can be made. The attenuation model better explains the quantitative data. Relief of early transcription termination in bacteria generally results in transcriptional differences of 5–20-fold. As mentioned, this is the order of magnitude of the transcriptional differences observed in eukaryotic systems upon induction of gene expression.

A certain amount of nonspecific transcription readthrough is predicted by the model. Transcription readthrough of normally attenuated sequences can account for the very extensive overlapping of sets of transcribed sequences between different tissues and embryonic stages [21,26,27]. Transcription readthrough is also a cogent explanation for the very low frequency of occurrence of most of the hnRNA and pre-mRNA [21,26,27]. Therefore, most of the sequence complexity of heterogeneous nuclear RNA observed in eukaryotes may be a consequence of such nonspecific transcription readthrough.

Early transcription termination (attenuation) would normally result in the formation of small abortive transcripts in the nucleus. If the termination complex were slowly released from the DNA template, a substantial part of this RNA might associate with the chromatin. The existence of low molecular weight 'cRNA' [44] hybridizing to a relatively large portion of the genome has been observed in a number of laboratories. It is possible the 'cRNA' is composed, to a large extent, of the products of early transcription termination rather than random breakdown products of hnRNA as suggested [45]. At times of active gene expression one might expect a diminution of such abortive transcription and a significant increase in the average length of the primary transcripts.

If the proposed model is correct, one would also expect to observe changes in the efficiency of attenuation by incorporation into RNA of base analogs which alter normal hydrogen bonding and consequently alter secondary structure. An increased efficiency of attenuation would be indicated by a decrease in the formation of large hnRNA molecules without a corresponding decrease in the formation of low molecular weight, attenuated RNA sequences. This is exactly what is found when the RNA synthesis inhibitor (an adenosine analog), 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) is administered to Hela cells [46]. The synthesis of RNA in the size-range of 140–740 bases catalyzed by RNA polymerase II is relatively unaffected by DRB while hnRNA synthesis is rapidly reduced by 66%. The 'resistant' hnRNA synthesis may represent a continued, though reduced, transcription readthrough. DRB also inhibits > 95% of mRNA synthesis. This implies, in addition, that the antiattenuator molecules are no longer effective in selectively preventing the attenuation (abortive transcription) of active genes. The site of action of DRB is evidently distal to the initiation point of RNA synthesis [46].

Antiattenuation during induction of gene expression should cause changes in the frequency distribution of classes of nuclear RNA. Such changes have been observed [29]. An increase in the length of specific pre-mRNA has, in fact, been observed during estrogen enhancement of vitellogenin sequence transcription in chick liver [18]; a dramatic increase in the length of the primary transcript formed in vitro using a homologous RNA polymerase II coincided with a 13-fold increase in vitellogenin transcript accumulation. More experiments of this genre should be performed to determine whether transcript elongation is a general phenomenon during eukaryotic gene activation.

The hypothesis presented here, modeled after attenuation of gene transcription in bacteria, does not postulate any cumbersome mutational events. The points emphasized in the model can be tested experimentally. The model has heuristic value in that it would account for and integrate much hitherto unexplained experimental data on gene transcription.

Acknowledgments

I would like to thank W. Alan Anderson for helpful criticism of the manuscript and for the origin of the term 'antiattenuation'. This work was supported by a grant from the National Research Council of Canada.

References

[1] Lee, F. and Yanofsky, Y. (1977) Proc. Natl. Acad. Sci. USA 74, 4365–4369.

[2] Goldberg, S., Schwartz, H. and Darnell, J. E. (1977) Proc. Natl. Acad. Sci. USA 74, 4520–4523.
[3] Shearer, R. W. and Smuckler, E. (1972) Cancer Res. 32, 339–343.
[4] O'Farrell, P., Cordell, B., Valenzuela, P., Rutter, J. and Goodman, H. (1978) Nature 274, 438–445.
[5] Tilghman, S. M., Curtis, P., Tiemeier, D., Leder, P. and Weissmann, C. (1978) Proc. Natl. Acad. Sci. USA 75, 1309–1313.
[6] Proudfoot, N. J. and Brownlee, G. (1976) Nature 263, 211–215.
[7] Baralle, F. E. (1977) Nature 267, 279–282.
[8] Crick, F. (1971) Nature 234, 25–28.
[9] Britten, R. J. and Davidson, E. H. (1969) Science 165, 349–356.
[10] Georgiev, G. P. (1969) J. Theor. Biol. 25, 473–484.
[11] Salser, W. (1978) Cold Spring Harbor Symp. Quant. Biol. 42, 985–1002.
[12] Paul, J. and Gilmour, R. S. (1968) J. Mol. Biol. 34, 305–314.
[13] Stein, G. S., Spelsberg, T. C. and Kleinsmith, L. J. (1974) Science 183, 817–824.
[14] Spelsberg, T. C., Mitchell, W. M., Chytil, F., Wilson, E. and O'Malley, B. (1973) Biochim. Biophys. Acta 312, 765–777.
[15] Zazloff, M. and Felsenfeld, G. (1977) Biochemistry 16, 5135–5142.
[16] Chambon, P. (1978) Cold Spring Harbor Symp. Quant. Biol. 42, 1209–1234.
[17] Towle, H. C., Tsai, M. J., Tsai, S. Y. and O'Malley, B. W. (1977) J. Biol. Chem. 252, 2396–2401.
[18] Dierks-Ventling, C. (1978) Nucl. Acids Res. 5, 3929–3943.
[19] Palmiter, R. D., Mulvihill, E., McKnight, G. and Senear, A. (1978) Cold Spring Harbor Symp. Quant. Biol. 42, 639–647.
[20] Orkin, S. and Swerdlow, P. (1977) Proc. Natl. Acad. Sci. USA 74, 2475–2479.
[21] Kleene, L. and Humphreys, T. (1977) Cell 12, 143–155.
[22] Melli, M., Spinelli, G. and Aranold, E. (1977) Cell 12, 167–174.
[23] Jansing, R. L., Stein, J. and Stein, G. S. (1977) Proc. Natl. Acad. Sci. USA 74, 173–177.
[24] Ryffel, G. and McCarthy, B. (1975) Biochemistry 14, 1379–1384.
[25] Hastie, N. and Bishop, J. O. (1976) Cell 9, 761–774.
[26] Young, B., Birnie, G. and Paul, J. (1976) Biochemistry 15, 2823–2829.
[27] Chikaraishi, D., Deeb, S. and Sueoka, N. (1978) Cell 13, 111–120.
[28] Hough, B., Smith, M., Britten, R. and Davidson, E. H. (1975) Cell 5, 291–299.
[29] Mauron, A. and Spohr, G. (1978) Nucl. Acids Res. 5, 3013–3032.
[30] Kasai, T. (1974) Nature 249, 523–527.
[31] Stauffer, G. V., Zurawski, G. and Yanofsky, C. (1978) Proc. Natl. Acad. Sci. USA 75, 4833–4837.
[32] Barnes, W. M. (1978) Proc. Natl. Acad. Sci. USA 75, 4281–4285.
[33] Zurawski, G., Brown, K., Killingly, D. and Yanofsky, C. (1978) Proc. Natl. Acad. Sci. USA 75, 4271–4275.
[34] Tsai, S., Tsai, M., Schwartz, R., Kalimi, M., Clark, J. and O'Malley, B. (1975) Proc. Natl. Acad. Sci. USA 72, 4228–4233.
[35] Goldstein, L., Wise, G. and Ko, C. (1977) J. Cell Biol. 73, 322–331.
[36] Frenster, J. H. (1976) Cancer Res. 36, 3394–3398.
[37] Krause, M. O. and Ringuette, M. (1977) Biochem. Biophys. Res. Commun. 76, 796–803.
[38] Deshpande, A., Jakowlew, S., Arnold, H., Crawford, P. and Siddiqui, M. (1977) J. Biol. Chem. 252, 6521–6527.
[39] Kanehisa, T., Kitazume, Y., Kazuyoshi, I. and Tanaka, Y. (1977) Biochim. Biophys. Acta 475, 501–513.
[40] Kano, Y., Komatsu, H., Nakawoin, K. and Fujiwara, Y. (1978) Exp. Cell Res. 115, 444–448.
[41] Jelinek, W. and Leinwand, L. (1978) Cell 15, 205–214.
[42] Shine, J. and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342–1346.
[43] Kozak, M. (1978) Cell 15, 1109–1123.
[44] Holmes, D., Mayfield, J. and Bonner, J. (1974) Biochemistry 13, 849–855.
[45] Weinberg, R. A. (1973) Ann. Rev. Biochem. 42, 329–354.
[46] Tamm, I. (1977) Proc. Natl. Acad. Sci. USA 74, 5011–5015.
[47] Wold, B., Klein, W., Hough-Evans, B., Britten, R. and Davidson, E. H. (1978) Cell 14, 941–950.