Mechanisms of Transcriptional Timing in *Drosophila*

CARL S. THUMMEL

ALTHOUGH IT IS CLEAR THAT TRANSCRIPTIONAL ACTIVATION and repression are important in dictating temporal patterns of gene expression, the molecular mechanisms that control the timing of transcription have remained largely elusive. Our lack of understanding stems from the scarcity of biological regulatory systems that have a discrete initiation event to provide a reference point for synchronizing the system and timing subsequent regulatory interactions. Studies of gene expression in viruses and the fruit fly, *Drosophila melanogaster*, have revealed that the physical arrangement and lengths of transcription units can play an important role in controlling their timing of expression.

In viruses, infection of a host cell initiates a cascade of viral gene expression that results in viral DNA replication and virion assembly. Some bacteriophages, such as T7, contain genomes in which the linear arrangement of genes determines their order of expression after infection. These genomes are injected into the host cell in a fixed orientation such that the early regulatory genes are introduced first, and the structural genes required for phage assembly are introduced last. It is the physical presentation of the viral template DNA that allows RNA polymerase to transcribe each set of genes in order, thereby determining the temporal order of viral gene expression (*1*). Similarly, transcription of the long late operon (~25 kb) encoding the structural components of the P22 and λ virions results in the sequential appearance of its encoded proteins, reflecting the positions of their coding regions along the DNA (*2*).

This observation in λ led to the formulation of a model, which proposed that biological time could be measured by transcription, taking advantage of its constant relatively slow rate (~2 kb/min for *Escherichia coli* RNA polymerase) (*3*). According to this model, delays in gene expression are determined by the length of time required to transcribe the gene itself, or another gene that encodes a critical transcriptional activator (*3*). On repressing transcriptional initiation, the length of a transcription unit can also establish a delay before there is a decrease in the number of newly synthesized messenger RNAs.

A similar model for gene length functioning as a delay timer has been invoked in *Drosophila* to propose a function for the unusual length of the transcription units associated with the homeotic genes (*4*). These genes, such as *Ultrabithorax* (*Ubx*) and *Antennapedia* (*Antp*), are responsible for establishing the identities of individual segments and thus play a relatively late role in embryonic pattern formation (*5*). The homeotic genes span long regions of DNA (77 and 100 kb, respectively, for *Ubx* and *Antp*) because they contain unusually long introns (as long as 50 and 60 kb, respectively, for *Ubx* and *Antp*) (*4, 6*). The delay inherent in the transcription of these genes was proposed to contribute to their temporal regulation, consistent with the observation that they function later in development than shorter segmentation genes (*4*).

Howard Hughes Medical Institute, Eccles Institute of Human Genetics, University of Utah, Salt Lake City, UT 84112.

Studies of *Ubx* transcription have demonstrated that gene length can function as a delay timer in *Drosophila* (*13*). Elongating nascent *Ubx* transcripts can be detected by in situ hybridization in precisely staged early *Drosophila* embryos. When the promoter is first activated, at the beginning of cell cycle 14, *Ubx* transcripts are elongated at a rate of ~1.4 kb/min, similar to the RNA polymerase II elongation rate determined in other higher organisms. This leads to an approximate 1-hour delay in the appearance of mature *Ubx* mRNA. Interestingly, *Ubx* transcription is aborted with the onset of mitosis and must reinitiate at the beginning of the next cell cycle. Because the lengths of the cell cycles in early *Drosophila* embryos are similar to the amount of time required to complete the first rounds of *Ubx* transcription, the length of the cell cycle determines how much *Ubx* product can accumulate. Thus, during cell cycle 14, *Ubx* mRNA can only accumulate for ~20 min before the next mitotic wave disrupts the transcription complex. At later times in development the embryo is subdivided into increasingly restricted mitotic domains, within which all cells undergo synchronous mitoses. It seems likely that the lengths of the cell cycles within each mitotic domain will contribute to determining how much *Ubx* product accumulates in specific cell types.

It is also in *Drosophila* that one of the most striking demonstrations of precise, temporally coordinated gene expression can be visualized as waves of puffing activity in the salivary gland polytene chromosomes (*7*). This puffing response is triggered by a pulse of the steroid hormone ecdysone at the onset of metamorphosis. Ecdysone directly activates six so-called early puffs, which appear to encode regulatory proteins that both repress their own expression and induce more than 100 late puffs over a 10-hour period (*8*). The sequential appearance of the late puff products is thought to direct the tissue-specific changes associated with the early stages of metamorphosis. This puffing response is reminiscent of viral life cycles, in which the initial expression of a few regulatory genes directs the subsequent expression of a much larger set of late structural genes.

Molecular characterization of the ecdysone regulatory hierarchy has provided further evidence for gene length functioning as a delay timer. *E74*, the gene responsible for the early puff at position 74EF, is composed of two nested transcription units, designated *E74A* and *E74B*, which have unique promoters but share a common 3′ end. On activation of its promoter by ecdysone, *E74A* is transcribed at a rate of ~1.1 kb/min, similar to that of *Ubx*. This rate, combined with the 60-kb length of the *E74A* transcription unit, accounts for most of the 1-hour delay in the appearance of its mRNA after ecdysone addition (*9*). In a similar manner, *E74B* mRNA begins to accumulate between 15 and 30 min after ecdysone addition, consistent with the 20-kb length of its primary transcript (*10*). Thus, as with T7 bacteriophage, the structure of the *E74* gene dictates an invariant order in the appearance of its products, whereby *E74B* expression will always precede that of *E74A* in response to ecdysone. Furthermore, whereas in most genes the length and sequences of introns diverge at a faster rate than protein coding regions, the length of the *E74A* unit is remarkably conserved in two distantly related species (*D. pseudoobscura* and *D. virilis*), consistent with its length playing an important role in *E74A* expression (*11*).

Although gene length contributes to the timing of *E74* transcription, the unique responses of the *E74A* and *E74B* promoters to different ecdysone concentrations is a significant factor in their temporal regulation (*10*). The structure of the *E74* gene thus establishes only one level of temporal control, an invariant parameter that must be further modulated by trans-acting regulators to provide the necessary biological flexibility during development. It should be possible to examine the contribution of *E74A* transcript length to its temporal regulation by constructing a minigene that is missing the majority of the *E74A* introns but contains intact

cis-regulatory sequences and exons. Insertion of such a construct into a wild-type genome should lead to an earlier appearance of the *E74A* gene product and may exert visible effects on the timing of the puffing response and early stages of metamorphosis.

The mechanisms that coordinate the timing of *E74* transcription appear to be used by other early genes in the ecdysone regulatory hierarchy. Two other early puff loci have been defined at the molecular level, and both correspond to unusually long and complex transcription units. The ecdysone-inducible transcription units within these puffs derive from nested promoters, encompass from 20 to 100 kb of DNA, and, like *E74*, encode potential transcriptional regulators (*12*). It seems likely that the lengths of these early transcription units define their order of appearance in response to ecdysone. Unique combinations of early gene products could then act on each set of late genes to coordinate the timing of their induction and direct the orderly progression of gene expression that programs the initial steps in metamorphosis.

Given the evidence that gene length can function as a delay timer, it is interesting that one *Drosophila* gene has been identified that is longer than the homeotic and early ecdysone-inducible genes. This gene, *dunce*, encodes adenosine 3′,5′-monophosphate (cAMP) phosphodiesterase and spans more than five bands in the polytene chromosomes (>148 kb, the exact length is not known). Although the absence of *dunce* mRNA in early embryos is consistent with the unusual length of its transcription unit, it is not known if the delay in *dunce* expression is of functional significance during development (*14*).

Transcriptional timing mechanisms similar to those seen in *Drosophila* are likely to occur in other higher organisms. The remarkable length of the gene encoding the human dystrophin protein, more than two megabases, could be a critical component in determining the amounts of dystrophin protein that can accumulate in mitotically active cells (*15*). Similarly, it may not be coincidence that the shortest transcription unit in the adenovirus genome is the first to be expressed, the immediate-early E1a gene (1.1 kb). The other early genes range up to 4.6 kb in length, with the exception of the 23-kb early transcription unit that encodes the viral DNA polymerase. Because the onset of viral DNA replication defines the switch from early to late gene expression, the unusual length of this early gene may contribute to the timing of this transition. Similarly, the late genes are all transcribed from a long operon that encompasses most of the adenoviral genome, spanning ~26 kb (*16*).

Other eukaryotic systems are being studied that afford an ideal opportunity to characterize the mechanisms that regulate transcriptional timing. The addition of growth factors to serum-starved tissue culture cells results in the synchronous induction of a large set of immediate-early genes. Many of these are protooncogenes, such as c-*fos*, c-*myc*, and c-*jun*, that encode well-characterized transcriptional regulatory proteins. The same set of immediate-early protooncogenes are also induced in fully differentiated neuronal cells by a variety of stimuli, including neurotransmitters and electrical excitation (*17*). The relatively short lengths of these proto-oncogenes are consistent with their immediate appearance upon induction. It will be interesting to learn whether the transcriptional timing mechanisms used in these signal-transduction pathways will be similar to those that contribute to the timing of gene expression in *Drosophila*.

REFERENCES AND NOTES

1. C.-C. Pao and J. F. Speyer, *J. Virol.* **11**, 1024 (1973); W. T. McAllister and C. Morris, *J. Mol. Biol.* **153**, 527 (1981).
2. P. N. Ray and M. L. Pearson, *Nature* **253**, 647 (1975); S. Casjens and M. B. Adams, *J. Virol.* **53**, 185 (1985).
3. J. D. Watson, *Molecular Biology of the Gene* (Benjamin, Menlo Park, CA, ed. 2, 1970), pp. 528–529.
4. K. Kornfeld *et al.*, *Genes Dev.* **3**, 243 (1989); D. Gubb, *Dev. Genet.* **7**, 119 (1986).
5. M. Akam, *Development* **101**, 1 (1987); P. W. Ingham, *Nature* **335**, 25 (1988).
6. M. B. O'Connor, R. Binari, L. A. Perkins, W. Bender, *EMBO J.* **7**, 435 (1988); S. Schneuwly, A. Kuroiwa, P. Baumgartner, W. J. Gehring, *ibid.* **5**, 733 (1986).
7. H. J. Becker, *Chromosoma* **10**, 654 (1959); M. Ashburner, *ibid.* **38**, 255 (1972).
8. M. Ashburner, C. Chihara, P. Meltzer, G. Richards, *Cold Spring Harbor Symp. Quant. Biol.* **38**, 655 (1974).
9. C. S. Thummel, K. C. Burtis, D. S. Hogness, *Cell* **61**, 101 (1990).
10. F. D. Karim and C. S. Thummel, *Genes Dev.* **5**, 1067 (1991).
11. C. W. Jones, M. W. Dalton, L. H. Townley, *Genetics* **127**, 535 (1991).
12. K. B. Burtis, C. S. Thummel, C. W. Jones, F. D. Karim, D. S. Hogness, *Cell* **61**, 85 (1990); R. Janknecht, W. Taube, H.-J. Lüdecke, O. Pongs, *Nucleic Acids Res.* **17**, 4455 (1989); W. A. Segraves and D. S. Hogness, *Genes Dev.* **4**, 204 (1990); G. Feigl, M. Gram, O. Pongs, *Nucleic Acids Res.* **17**, 7167 (1989); P. R. DiBello, D. A. Withers, C. A. Bayer, J. W. Fristrom, G. M. Guild, *Genetics* **129**, 385 (1991).
13. A. W. Shermoen and P. H. O'Farrell, *Cell* 67 303 (1991).
14. R. L. Davis and N. Davidson, *Mol. Cell. Biol.* **6**, 1464 (1986); C. Chen, T. Malone, S. K. Beckendorf, R. L. Davis, *Nature* **329**, 721 (1987); R. L. Davis and B. Dauwalder, *Trends Genet.* **7**, 224 (1991).
15. M. Burmeister *et al.*, *Genomics* **2**, 189 (1988).
16. J. Tooze, Ed., *DNA Tumor Viruses* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, ed. 2, part 2, 1980).
17. K. Kelly, B. H. Cochran, C. D. Stiles, P. Leder, *Cell* **35**, 603 (1983); M. E. Greenberg and E. B. Ziff, *Nature* **311**, 433 (1984); B. H. Cochran, A. C. Reffel, C. D. Stiles, *Cell* **33**, 939 (1983); L. F. Lau and D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1182 (1987); M. Sheng and M. E. Greenberg, *Neuron* **4**, 477 (1990); J. I. Morgan and T. Curran, *Trends Neurosci.* **12**, 459 (1989); M. Dragunow, R. W. Currie, R. L. M. Faull, H. A. Robertson, K. Jansen, *Neurosci. Biobehav. Rev.* **13**, 301 (1989).

18. I thank D. Hogness for stimulating discussions and his contributions to the ideas presented here; F. Karim, L. Boyd, S. Sakonju, G. Herrick, and P. O'Farrell for critical comments on the manuscript; and P. O'Farrell for communicating results before publication. Our research is supported by the Howard Hughes Medical Institute and NIH.