Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 9384–9388, September 1996
Colloquium Paper

This paper was presented at a colloquium entitled "Biology of Developmental Transcription Control," organized by Eric H. Davidson, Roy J. Britten, and Gary Felsenfeld, held October 26–28, 1995, at the National Academy of Sciences in Irvine, CA.

Chromatin structure and gene expression

GARY FELSENFELD, JOAN BOYES, JAY CHUNG*, DAVID CLARK†, AND VASILY STUDITSKY

National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0540

ABSTRACT **It is now well understood that chromatin structure is perturbed in the neighborhood of expressed genes. This is most obvious in the neighborhood of promoters and enhancers, where hypersensitivity to nucleases marks sites that no longer carry canonical nucleosomes, and to which transcription factors bind. To study the relationship between transcription factor binding and the generation of these hypersensitive regions, we mutated individual cis-acting regulatory elements within the enhancer that lies between the chicken β- and ε-globin genes. Constructions carrying the mutant enhancer were introduced by stable transformation into an avian erythroid cell line. We observed that weakening the enhancer resulted in creation of two classes of site: those still completely accessible to nuclease attack and those that were completely blocked. This all-or-none behavior suggests a mechanism by which chromatin structure can act to sharpen the response of developmental systems to changing concentrations of regulatory factors. Another problem raised by chromatin structure concerns the establishment of boundaries between active and inactive chromatin domains. We have identified a DNA element at the 5′ end of the chicken β-globin locus, near such a boundary, that has the properties of an insulator; in test constructions, it blocks the action of an enhancer on a promoter when it is placed between them. We describe the properties and partial dissection of this sequence. A third problem is posed by the continued presence of nucleosomes on transcribed genes, which might prevent the passage of RNA polymerase. We show, however, that a prokaryotic polymerase can transcribe through a histone octamer on a simple chromatin template. The analysis of this process reveals that an octamer is capable of transferring from a position in front of the polymerase to one behind, without ever losing its attachment to the DNA.**

DNA is packaged as chromatin within the nuclei of eukaryotes, and in the neighborhood of genes, this compact structure must be disrupted when the genes are transcribed by RNA polymerase II. We have a reasonably good idea of what happens at the lowest levels of chromatin organization; nucleosomes located at nearby promoters and enhancers are disrupted or displaced, giving rise to short regions (hypersensitive sites, HSs) that are unusually sensitive to nucleases and chemical probes. However, the body of the gene continues to a considerable extent to be packaged in nucleosomes, and there are manifestations of higher order structure as well (ref. 1; Fig. 1).

Two questions are raised by this structure: How is an active chromatin structure established so that RNA polymerase II can initiate transcription, and, once initiated, how does the polymerase manage to traverse the typical nucleosome-covered gene? Work in our laboratory has addressed these two issues.

What Is Required to Establish a Hypersensitive Domain?

To study the relationship between chromatin structure and gene activation, we have chosen the chicken β-globin gene family as a typical group of developmentally regulated, transcriptionally active genes. There are four genes in the cluster (Fig. 2), two of which, ε and ρ, are embryonic genes expressed in the primitive lineage; the other two, β^H and the adult beta globin gene β^A, are expressed in the definitive lineage. Much work in our own and other laboratories has defined the promoters of these genes as well as more distant regulatory elements. Here we focus attention on the β^A-globin gene, which is controlled by nearby upstream elements and also by a strong enhancer (3, 4) that lies at the 3′ end of the gene. This enhancer functions bidirectionally, activating as well the downstream ε-globin gene in primitive lineage cells. The region containing the enhancer is strongly hypersensitive to nucleases, and the pattern of hypersensitivity is consistent with the absence of a nucleosome. Within the enhancer, there are five binding sites for four distinct regulatory factors, as shown in the lower part of Fig. 2.

In earlier studies, we used transgenic mice to examine the role of the enhancer in stimulating expression from the β^A-globin gene (5); a separate issue is the mechanism by which the active chromatin structure is generated. In particular, we were interested in knowing whether the enhancer was an autonomous element that would be hypersensitive in the absence of the promoter. Therefore, we made transgenic mice carrying constructions in which either the β^A-globin promoter or enhancer was deleted, and we determined the DNase I sensitivity of the remaining element (6). The pattern of sensitivity in a series of mouse lines was consistent with a model in which enhancer hypersensitivity depended on the presence nearby either of the β^A-globin promoter or of some other promoter that was near the point of insertion into the mouse genome. Hypersensitivity was always observed at both promoter and enhancer when both were present, and it was never observed at the promoter in the absence of the enhancer. Thus, the promoter and enhancer of this gene appear to interact to generate the HSs seen at each.

We next wished to ask what governed the hypersensitivity of the enhancer. Mutagenesis studies had shown that the sites which bind the erythroid factors NFE2 and GATA-1 are most important for activating transcription (7). For this purpose (2),

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.

Abbreviation: HS, hypersensitive site.

*Present address: Molecular Hematology Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD.

†Present address: Laboratory of Cellular and Developmental Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD.

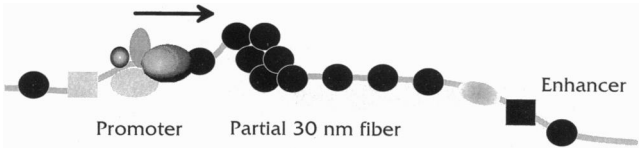

FIG. 1. Schematic diagram of chromatin structure at a transcriptionally active gene. The dark circles are nucleosomes; other symbols represent components of the transcription complex at the promoter (arrow), as well as more distant transcriptional activators. The nucleosomes are shown partially folded into a higher order structure.

we made use of the same mutants that had been employed (7) in studying effects on transcription, but we examined effects on hypersensitivity. Constructions carrying the β^A-globin promoter, the gene, and the enhancer, mutated to remove each of the factor binding sites in turn and in combination, were stably introduced into the avian erythroblastosis virus-transformed erythroid precursor cell line 6C2. Mutation of binding sequences for the factors NFE2 and GATA-1 reduced the accessibility to enzymic probes, and this diminution increased additively as these sites were successively destroyed. There are two general mechanisms that could give rise to this behavior (Fig. 3). Perhaps as GATA-1 or NFE2 sites are removed, the HS becomes increasingly inaccessible sterically so that digestion is slower but will ultimately reach completion. Alternatively, there could be two populations, one with HSs completely blocked and the other with HSs completely accessible.

It is impossible to distinguish these models with DNase I, the normal probe for hypersensitivity, because its digestion endpoint is only reached when the DNA is completely fragmented. Therefore, we carried out the digestion with one of two restriction enzymes that cut within the region when it is hypersensitive. We compared both the kinetics and endpoint of digestion with the internal control provided by the endogenous site. We found, for example, that an HS with both GATA-1 binding sequences mutated is digested to a plateau value considerably lower than that for the endogenous HS, indicating that only about 40% of the sites are accessible. If the data are plotted with normalized plateaus for the endogenous and mutant sites, the rates of digestion for accessible sites are seen to be the same. Thus, damaging the hypersensitive region by reducing the number of bound transcription factors reduces the probability that a site will be hypersensitive, but any site that is active is fully accessible. These results are consistent with, and may well explain, earlier results that show an all-or-none effect of enhancers on expression (8, 9). They also suggest a way in which chromatin organization can serve to sharpen the transcriptional response of a gene to a slowly varying factor concentration.

How Is a Transcriptionally Active Chromatin Domain Demarcated from Surrounding Inactive Regions?

The entire β-globin locus is marked by a general sensitivity to nucleases that is about 3- to 5-fold higher than that of surrounding DNA sequences. Work in our laboratory some years ago (10) had identified a strong, constitutive DNase I HS (5′HS4) present in all tissues, about 20 kb upstream of the 5′ end of the ρ-globin gene, which we thought might mark the 5′ boundary of the locus. At that time, studies by Schedl and coworkers (11) had just characterized a *Drosophila* element, *scs*, which appeared to mark a chromatin structural boundary and which also served as an insulating element, protecting a mini-*white* reporter gene against position effects when integrated into the *Drosophila* genome. Therefore, we asked whether the β-globin element had similar properties. For the purposes of our assay, we constructed (Fig. 4*A*; ref. 12) a reporter carrying a gene for neomycin (G418) resistance, coupled to the promoter of the human γ-globin gene and a strong enhancer, the mouse β-globin locus control element (5′HS2). When this is stably integrated into the human erythroleukemia cell line K562, selection for neomycin resistance results in the appearance of a large number of colonies. A 1.2-kb fragment containing the chicken 5′HS4 was then tested for insulating activity by inserting it on both sides of the promoter/gene pair (Fig. 4*B*) so that it was interposed between enhancer and promoter even at sites of tandem integration.

The results of this experiment are shown in Fig. 5*A*. There was a marked decrease in the number of resistant colonies relative to a control containing an equivalent length of λ phage DNA; the effect was even greater when two copies of the element were used on each side. In a variant of this experiment, two markers, one of which was surrounded by the chicken element, were introduced on the same DNA fragment (Fig. 5*B*). The ratio of G418-resistant colonies to hygromycin-resistant colonies decreased markedly when one or two copies of the element were introduced. The results were again consistent with a strong activity of the chicken element in blocking action of the enhancer on the promoter. In an experiment exactly parallel to that by Kellum and Schedl (11) for the *scs* element, we surrounded the *white* minigene on each side with two copies of the 1.2-kb insulator fragment (12) and found that the insulator also protected this test gene from position effects in *Drosophila*.

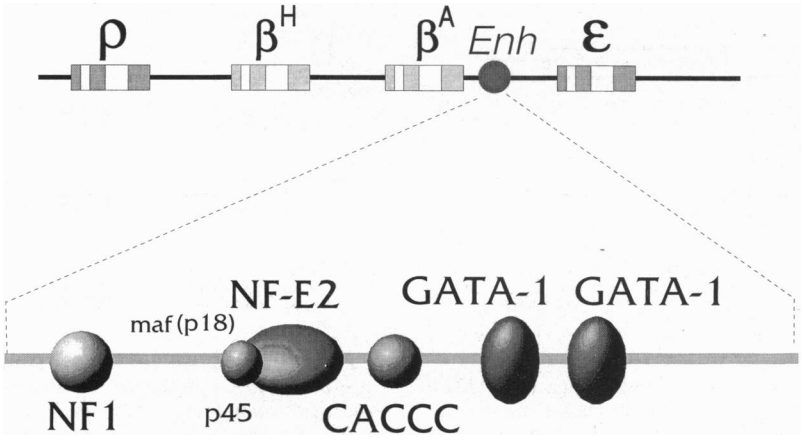

FIG. 2. The chicken β-globin gene cluster, showing the position of the HS between the β- and ε-globin genes, which is a strong enhancer. Below it is the detailed structure of this enhancer, showing binding sites for transcription factors, including the erythroid-specific factors GATA-1 and NFE-2 (see ref. 2).

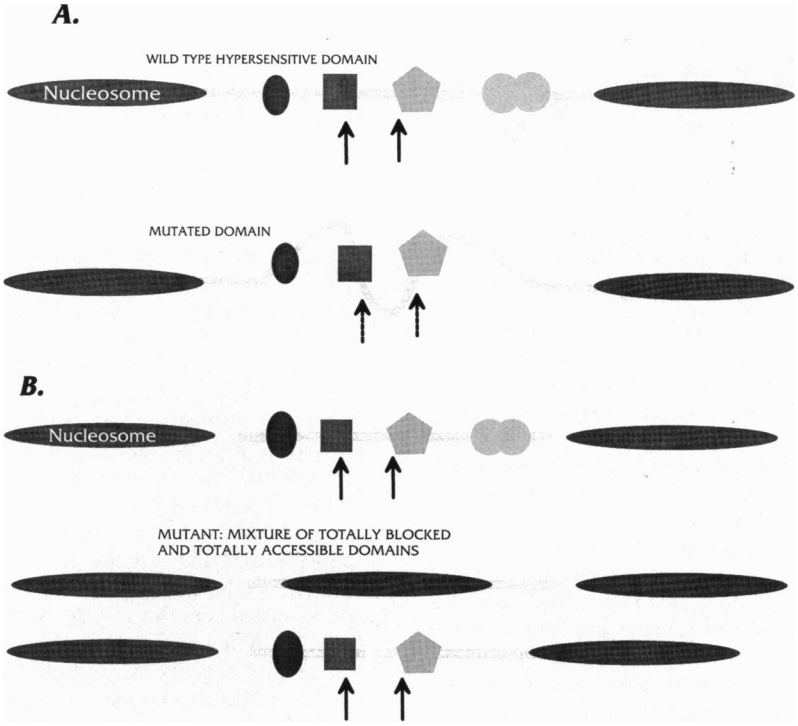

FIG. 3. Two possible mechanisms for generating a "weakened" HS. (*A*) Mutation of binding sites within the HS results in a uniform population of less accessible sites. (*B*) Mutation results in two populations, one fully accessible and the other totally inaccessible.

Does the chicken insulator play any role in establishment of domain boundaries? The boundaries of the active β-globin domain recently have been determined with considerable precision (13). The pattern of general sensitivity to DNase I mentioned above shows a sharp decrease as one moves 5′ across the site of the insulator, as does the level of histone acetylation, another indicator of transcriptionally active chromatin. There is thus a striking correlation between the position of the insulator and the end of the transcriptionally active domain. Other experiments must be devised to determine whether the insulator actually participates in creation of the boundary.

We do not know how the β-globin element works as an insulator. The insulation activity of the *gypsy* element in *Drosophila* depends in part on a protein, the product of the gene *suppressor of hairy wing*, which binds to motifs within the element (14). Another element with insulating properties, *scs′*, has been shown to bind the protein BEAF32 (15). It seems likely that the β-globin insulator also requires participation of DNA binding proteins. We have, therefore, been using deletion analysis and mutagenesis in attempts to narrow down the site responsible for insulation. These studies show that a

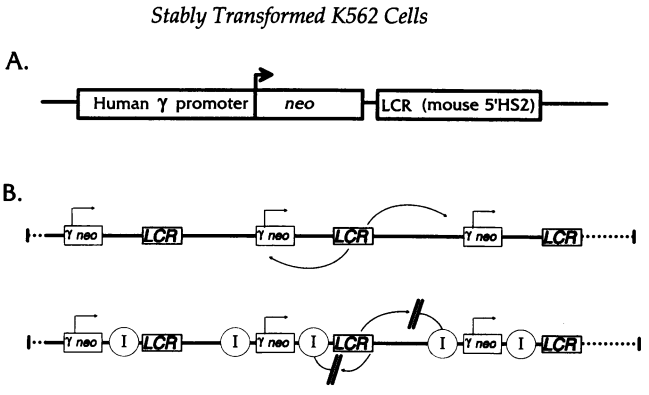

FIG. 4. DNA constructions used for testing insulating activity (12). (*A*) The neomycin resistance gene is coupled to a promoter and strong enhancer/locus control element (LCR) as described in the text. (*B*) Control (*Upper*) and experiment in which tandem copies are shown after integration into K562 human erythroleukemia cells. In the experiment, a 1.2-kb DNA sequence element has been inserted on either side of the promoter/gene. I, insulator.

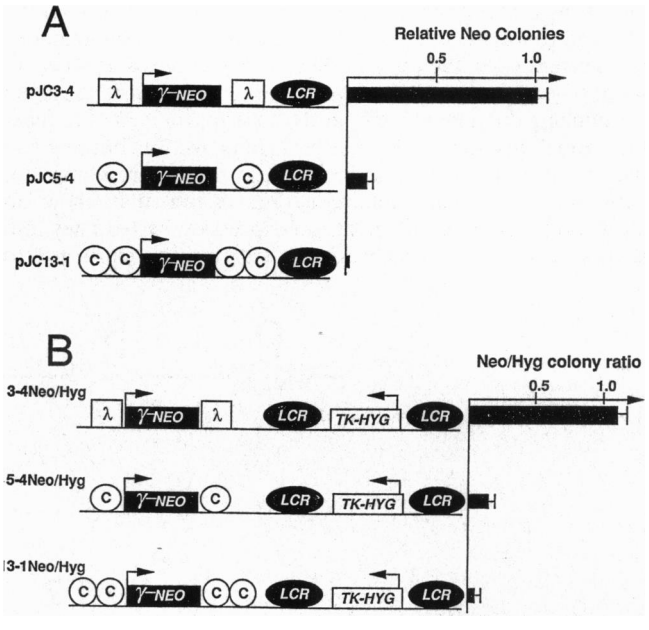

FIG. 5. (*A*) Resistance to G418 conferred by constructions made as shown in Fig. 4. In the first construction, DNA from λ phage DNA was used to maintain a similar spacing, and all other results for number of resistant colonies were normalized to this. C, insulator. (*B*) A similar experiment in which a second selectable but uninsulated marker for hygromycin resistance, coupled to a thymidine kinase promoter, was appended. The ratio of G418 to hygromycin resistance is shown (12).

250-bp sequence at the 5′ end of the 1.2-kb fragment discussed above retains much of the activity (unpublished data).

None of this sheds much light on the mechanism of insulator action. It has been suggested that insulators may serve as initiation sites for the directional formation of heterochromatin, a model in part inspired by studies of inactivation at telomeres and mating type loci in yeast, as well as by position effect phenomena in *Drosophila*. Another possibility is that insulators may function as "anchors" at each end of a domain. This would provide topological isolation of the regions they bound, so that enhancers outside the domain could not reach inside. A third class of models invokes "tracking" mechanisms in which a complex formed at a distant enhancer moves along the chromatin template until it reaches the promoter; the proposed function of the insulator is to derail this complex. None of these models satisfactorily explains the entire set of observations concerning insulators, but it is quite possible that not every insulator functions in the same way.

How Does RNA Polymerase Transcribe Through a Histone-Covered Template?

Although most of the attention has been focused on how chromatin domains are activated for transcription, it is equally important to understand how this "active" chromatin template is transcribed. There is evidence that genes transcribed by RNA polymerase II, such as those coding for β-globin, are packaged in nucleosomes in cells in which these genes are expressed (1). How does a polymerase manage to pass through such an obstacle? The following possibilities suggest themselves. (*i*) The histone octamer remains in place, perhaps by binding transiently to the nontranscribed strand. (*ii*) The octamer is displaced into solution. (*iii*) The octamer slides ahead of the polymerase. (*iv*) The octamer is displaced and recaptured.

We addressed these possibilities (16) by experiments in which a single nucleosome core particle was ligated into a plasmid so that it lay between an SP6 polymerase promoter upstream and transcription terminators downstream. The template was transcribed, and the position of the octamer was determined. The octamer could be recovered quantitatively bound to the plasmid, but it had moved more or less randomly, with some preference for the half of the plasmid 5′ of the promoter. These results eliminate all the possibilities except the fourth (above), but open the further issue of how the displacement occurs. In principle, the reaction might involve complete disruption of the histone–DNA interaction, with the octamer trapped for a time in the electrostatic field of the plasmid before recapture. A second possibility is that the octamer transfers by collision with some proximal or distal DNA sequence within the plasmid so that transfer occurs without the octamer ever letting go.

Recent work in our laboratory (17, 18) shows that the latter mechanism is the correct one. To reduce the possibilities for octamer movement, we reconstituted nucleosome core particles on short pieces of DNA carrying the SP6 promoter. The octamer typically occupies one of only a small number of preferred positions, and the positional isomers can be distinguished, separated, and characterized on a polyacrylamide gel. After transcription, the products can be analyzed similarly (Fig. 6). An octamer is typically displaced backward on the template by 40–80 bp. At low NTP concentrations, addition of competitor in excess does not result in transfer of octamer to the competitor, showing that during transcription the octamer is not displaced from the DNA to which it was originally bound. A schematic diagram of the mechanism that explains this behavior is shown in Fig. 7. The essential feature is that the octamer transfers from a position in front of the polymerase to one behind, without ever losing its attachment to DNA.

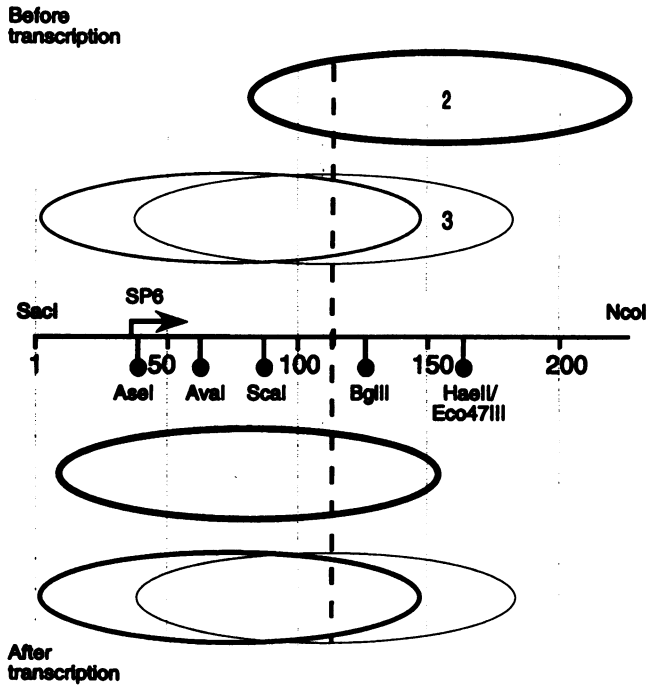

FIG. 6. Position of nucleosome on a defined sequence fragment before and after transcription. A 227-bp DNA fragment carrying an SP6 polymerase promoter was reconstituted with a single histone octamer, which occupied one of a small number of preferred positions. For each positional isomer, the position after transcription was also determined (17).

This model raises a further question: at what point in the progress of the polymerase does the transfer of the histone octamer occur? A different strategy was used to answer this question. A short nucleosome core template was employed

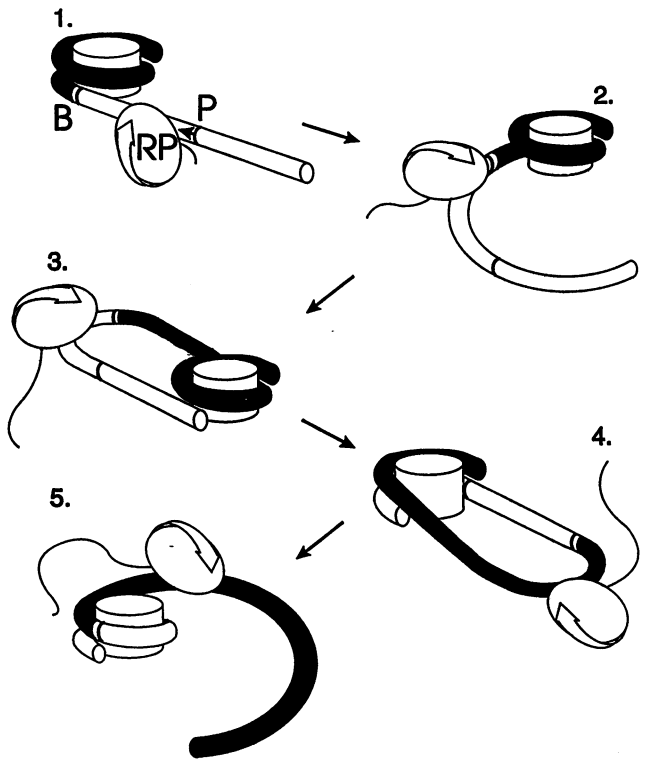

FIG. 7. Mechanism of transfer of an octamer from in front of the advancing polymerase to behind it.

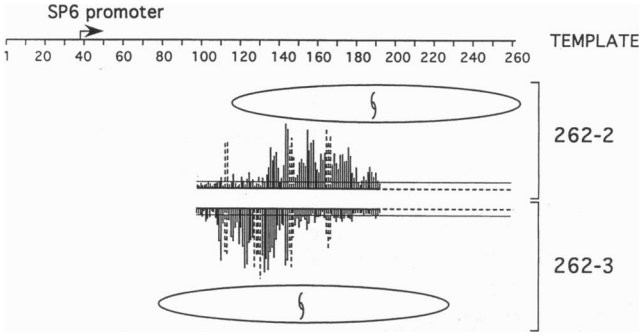

FIG. 8. Nucleosome core particles like those described in Fig. 4 were studied under conditions allowing synchronous transcription (see text). Pausing was observed during chain elongation, as shown here by solid vertical lines (the dashed lines represent pauses on naked DNA). The upper pattern was observed with a nucleosome positioned as shown near the top. The lower pattern was seen with the lower nucleosome position. Cessation of pausing occurs in each case near the nucleosome dyad axis; the DNA sequence is the same for both (18).

that contained a 16-nt "C-less" track to allow arrest at the initiation step and the labeling of the RNA 5′ terminus. Transcription was then permitted to proceed, and the RNA intermediates were examined. A pattern of pausing in elongation was observed (Fig. 8), which depended not on the DNA sequence but on the position of the histone octamer. Pausing began when the polymerase had advanced about 20 bp into the nucleosome and stopped near the nucleosome dyad axis. We conclude that the octamer is an impediment until the enzyme reaches the half-way point and that this is when the octamer transfers.

We also were interested in learning why the pausing occurs. One possibility is that the octamer simply provides a steric block to advance of the polymerase; another is that the closed loop intermediate generates constraints that alter the rate of advance. In further experiments (18), we used restriction enzymes to shorten the 5′ part of the template behind the polymerase after initiation. This has an effect on the elongation pattern consistent with a role for loop formation in pausing: the progress of the polymerase is slowed when the loop forms, and the polymerase tends to advance when the loop opens transiently.

These results show that a histone octamer is not an insuperable obstacle to the passage of an RNA polymerase. The octamer can be transferred around the enzyme without releasing its grip on DNA. The use of short DNA segments, of course, limits the final positions available to the octamer, providing an opportunity to study the mechanism. On longer templates, the octamer can travel to any vacant site on the DNA. Presumably, the likelihood of transfer to a given point on a long DNA is governed by the probability of "ring closure." In eukaryotic systems, there are no doubt auxiliary mechanisms that assist the passage of the polymerase. These may include histone acetylation and complexes that help to destabilize the octamer structure. Nonetheless, it seems likely that the intrinsic ability of the octamer to get out of the way will play a role in the eukaryotic transcription process.

Conclusion

The view that chromatin provides a mechanism for allowing DNA to fold in a compact form within the nucleus is certainly correct but in the past has tended to obscure the fact that chromatin and not DNA is the template for eukaryotic polymerases. Recent results make it clear that transcriptional mechanisms have learned not only to accommodate nucleosomes but also to take advantage of their properties for regulatory purposes (see for example ref. 19). To study the interactions of histones and the transcriptional mechanism, we chose the chicken globin gene family. Early studies were designed to determine the regulatory sites and corresponding factors that affected both the gene clusters and the individual genes. Now we are addressing the questions of how transcriptionally active chromatin is generated (6), how it is maintained (2), and how its limits are defined (12). At the same time, it is necessary to understand how a polymerase that binds to a promoter interacts with its chromatin environment, during both initiation and chain elongation. Even though the structure of the nucleosome is reasonably well understood, its biochemistry and that of the higher order structures it forms will surely turn out to be complex, and important to understanding eukaryotic gene regulation.

1. Felsenfeld, G. (1992) *Nature (London)* **355,** 219–224.
2. Boyes, J. & Felsenfeld, G. (1996) *EMBO J.* **15,** 2496–2507.
3. Hesse, J. E., Nickol, J. M., Lieber, M. R. & Felsenfeld, G. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4312–4316.
4. Choi, O.-R. & Engel, J. D. (1986) *Nature (London)* **323,** 731–734.
5. Reitman, M., Lee, E., Westphal, H. & Felsenfeld, G. (1990) *Nature (London)* **348,** 749–752.
6. Reitman, M., Lee, E., Westphal, H. & Felsenfeld, G. (1993) *Mol. Cell. Biol.* **13,** 3990–3998.
7. Reitman, M. & Felsenfeld, G. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 6267–6271.
8. Weintraub, H. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 5819–5823.
9. Walters, M. C., Fiering, S., Eidemiller, J., Magis, W., Groudine, M. & Martin, D. I. K. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7125–7129.
10. Reitman, M. & Felsenfeld, G. (1990) *Mol. Cell. Biol.* **10,** 2774–2786.
11. Kellum, R. & Schedl, P. (1991) *Cell* **64,** 941–950.
12. Chung, J. H., Whiteley, M. & Felsenfeld, G. (1993) *Cell* **74,** 505–514.
13. Hebbes, T. R., Clayton, A. L., Thorne, A. W. & Crane-Robinson, C. (1994) *EMBO J.* **13,** 1823–1830.
14. Corces, V. G. (1995) *Nature (London)* **376,** 462–463.
15. Zhao, K., Hart, C. M. & Laemmli, U. K. (1995) *Cell* **81,** 879–889.
16. Clark, D. J. & Felsenfeld, G. (1992) *Cell* **71,** 11–22.
17. Studitsky, V. M., Clark, D. J. & Felsenfeld, G. (1994) *Cell* **76,** 371–382.
18. Studitsky, V. M., Clark, D. J. & Felsenfeld, G. (1995) *Cell* **83,** 19–27.
19. Kingston, R., Bunker, C. A. & Imbalzano, A. N. (1996) *Genes Dev.* **10,** 905–920.