Discontinuous Mechanism of Transcription Elongation

Evgeny Nudler,* Alex Goldfarb, Mikhail Kashlev†

During transcription elongation, three flexibly connected parts of RNA polymerase of *Escherichia coli* advance along the template so that the front-end domain is followed by the catalytic site which in turn is followed by the RNA product binding site. The advancing enzyme was found to maintain the same conformation throughout extended segments of the transcribed region. However, when the polymerase traveled across certain DNA sites that seemed to briefly anchor the front-end domain, cyclic shifting of the three parts, accompanied by buildup and relief of internal strain, was observed. Thus, elongation proceeded in alternating laps of monotonous and inchworm-like movement with the flexible RNA polymerase configuration being subject to direct sequence control.

Elongation of transcription has been envisaged as a monotonous process in which each nucleotide addition is accompanied by a one–base pair (bp) translocation of RNA polymerase, which was considered an inflexible, rigid structure (*1*). Recent observations, however, suggest a substantial conformational plasticity of the ternary complex. First, deoxyribonuclease (DNase) I footprinting of RNA polymerase that has been walked in single-nucleotide steps through a short segment of DNA showed that the front end of the enzyme advances in a leap of several nucleotides (*2*). Other observations relate to the phenomenon of internal transcript cleavage induced by the GreB protein (*3, 4*). The cleavage removes the 3′-proximal segment of RNA, whereas the 5′ fragment remains in the ternary complex. Hence, it was hypothesized that the complex contains a 3′-proximal loose RNA-binding site, and a tight RNA-binding site upstream from the point of cleavage (Fig. 1A) (*5*). In a series of successive complexes, a 3′ segment of increasing length was cleaved and removed by GreB, reflecting gradual filling of the loose site with RNA. Thus, a two-stroke model of RNA chain growth emerged, in which the filling of the loose site alternates with the threading of the recently synthesized RNA through the tight site (*4, 6*). The coupling of the two-stroke mechanism of RNA synthesis with the leaplike translocation of RNA polymerase is the central postulate of the inchworm model of elongation that was first suggested and theoretically justified by Chamberlin (*7*). The principal unresolved question of the model is whether the inchworming is an obligatory, intrinsic feature of the elongation process or an incidental event that occurs only at specific DNA sequences. To address these issues we carried out systematic mapping of the three advancing parts of RNA polymerase in a succession of complexes halted along a long stretch of template.

We prepared defined elongation complexes using RNA polymerase that carries six His residues at the COOH-terminus of the β′ subunit (*8*). Through the His-tag, the enzyme adsorbs to the Ni^{2+}-chelating nitrilotriacetic acid (NTA)–agarose so that assays can be performed in solid phase. By centrifuging and resuspending the beads with adsorbed elongation complexes, it is possible for the polymerase to travel along the template in controlled steps (walking). Defined elongation complexes were prepared that spanned the segment +20 to +75 in the T7A1 transcription unit. The nomenclature used identifies elongation complexes by the length of the transcript; that is, EC_{70} designates the complex carrying 70-nucleotide (nt) RNA.

For each complex, several parameters were determined (Fig. 1A). The position of the catalytic site (C) followed directly from the length of the transcript. The front edge (F) was mapped by protection of the nontemplate DNA strand from 3′ to 5′ degradation with Exo III. The distance between the transcript tip and the front edge (C~F) was calculated from the two above parameters. The boundary between the tight and the loose RNA-binding sites (R) was mapped from the length of RNA remaining in the complex after cleavage with GreB. In some cases, the cleavage position was verified by identification of the released 3′ fragment (R~C). In most cases, halted complexes could resume elongation after the addition of transcription substrates (chasing). When a complex failed to resume elongation, it was classified as nonchaseable dead end according to the original designation by Chamberlin (*9*).

The relation between C and F within the segment +47 to +60 is shown in Fig. 2. The bottom panel shows RNA transcripts and is organized in three-lane sets corresponding to the intact halted complex (left lanes), the complex after treatment with Exo III (middle lanes), and the Exo III–treated complex after chasing (right lanes). The top panel presents the front edge mapping by Exo III digestion. The advancement between 47 and 53 is monotonous. Each nucleotide addition is accompanied by equivalent translocation of the front edge, so that the C~F distance maintains a constant value of 18 nt. By contrast, as the transcript grows from 53 to 57, a cycle of inchworming is observed. From 53 to 56 no advancement of the front end takes place. In fact, in EC_{56}, the front edge retreats by 2 bp as compared to EC_{54} to give the C~F value of 12 nt. EC_{56} is unstable and converts spontaneously into the nonchaseable dead-end complex. This conversion for EC_{56} is about 30%.

As the transcript grows by 1 nt from 56 to 57, the front end leaps ahead by 7 bp and the complex regains its original features, including the 18-nt C~F value. Further elongation from 57 to 60 is monotonous. Thus, within the segment +47 to +60 the elongating polymerase alternates between monotonous and inchworm-like modes of movement.

EC_{52} and EC_{56} were chosen for further analysis as representative complexes of the monotonous and the inchworming states, respectively. The effect on these complexes of different doses of GreB is shown in Fig. 3. The two complexes differ in GreB sensitivity. EC_{52} is resistant to moderate exposure to the factor (up to 1.5 GreB molecules per complex), whereas one-tenth the amount of GreB completely cleaves the transcript in EC_{56} remov-

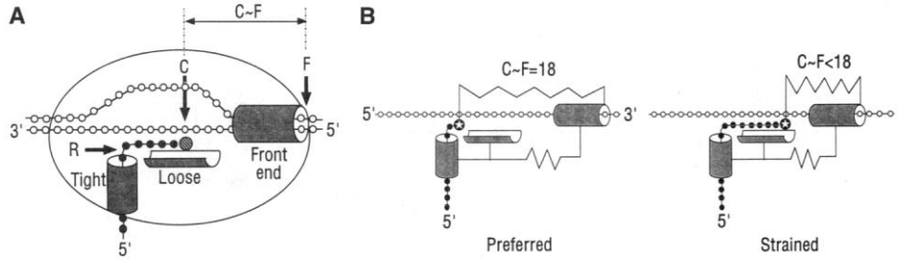

Fig. 1. States of the elongation complex. (**A**) Definition of components and dimensional parameters. (**B**) The two principal forms of the complex established in the present work.

Public Health Research Institute, 455 First Avenue, New York, NY 10016, USA.

*Is in the Graduate Program of the Institute of Molecular Genetics, Russian Academy of Sciences, Moscow, Russia.

†To whom correspondence should be addressed.

ing a 4-nt piece of RNA. The cleavage is accompanied by a forward shift of the front edge, so that the 18-nt C~F parameter is restored. The cleaved complex is fully chaseable. Thus, the cleavage effectively converts EC_{56} into EC_{52}. At higher doses of GreB, both the original EC_{52} and the EC_{52} derived from EC_{56} behave identically. RNA is cleaved in increments of 1 to 2 nt from the 3′ terminus, accompanied by retreating of the front edge. Thus, the attack with a high dose of GreB appears to push EC_{52} backward without major internal rearrangement (*10*). We interpret the two levels of GreB sensitivity as the reflection of two different phenomena: In one case, the cleavage occurs in the context of the static complex, in the other it requires backward translocation of the enzyme.

From these results, two types of elongation complexes can be distinguished (Fig. 1B). The GreB-resistant complex with C~F = 18 contains no RNA in the loose site and is characteristic of monotonous movement. Such a complex is relatively stable and, as shown below, is predominant during elongation. We designate this conformation of elongation complex "preferred." The GreB-sensitive complex with C~F < 18 carries RNA in its loose binding site, is unstable, and can spontaneously collapse into the dead-end configuration. The unstable complex behaves as if it contains internal tension, and we term it "strained."

The principal features of the preferred or strained conformations were found in all 25 complexes that we analyzed (Table 1). It can be concluded that within the segment +20 to +75, the "road map" of elongation includes two alternating monotonous (+30 to +53 and +57 to +75) and two inchworming (+23 to +27 and +54 to +56) stretches. Of all the complexes, only the last complex in each inchworming cycle (EC_{27} and EC_{56}, respectively) displayed a high level of conversion to dead end under standard conditions, as indicated by the asterisk.

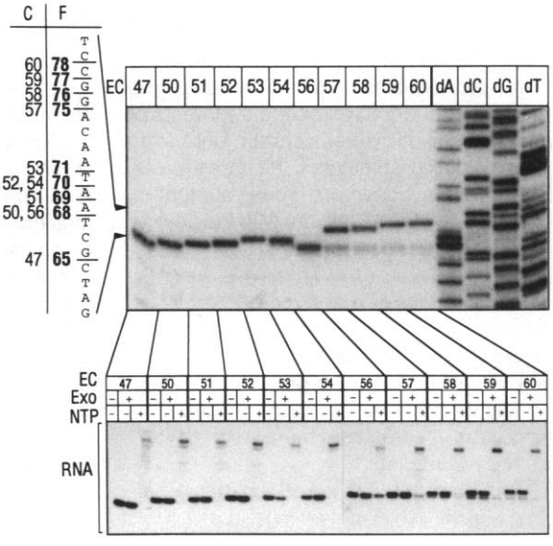

Fig. 2. Advancement of the elongation complex in the segment +47 to +60. RNA polymerase was walked through the segment in single-nucleotide steps (*14*), yielding defined elongation complexes from EC_{47} to EC_{60}. The top panel shows the protection of terminally ^{32}P-labeled nontemplate DNA strand from Exo III digestion. The bottom panel presents radiolabeled RNA in each complex analyzed either in the intact form or after treatment with Exo III before or after the chase with the four NTPs, as indicated (*15*). The table on the left lists the positions of the front end (F) in each complex which is identified by the length of the transcript (C).

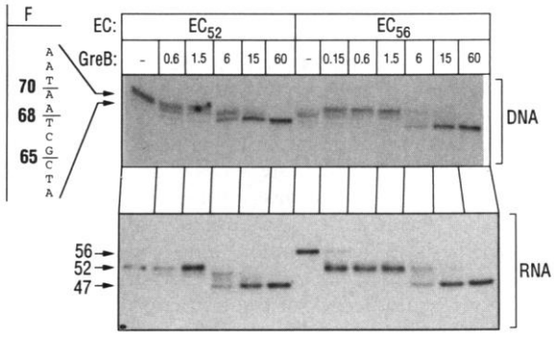

Fig. 3. RNA cleavage in elongation complexes EC_{52} and EC_{56}. The two complexes were exposed to increasing doses of GreB expressed in molar equivalents of the GreB protein per RNA polymerase (*16*). The top panel shows the protection of DNA from Exo III (*15*), and the bottom panel, the RNA in the complex. The table indicates the positions of the front end (F) on the DNA sequence.

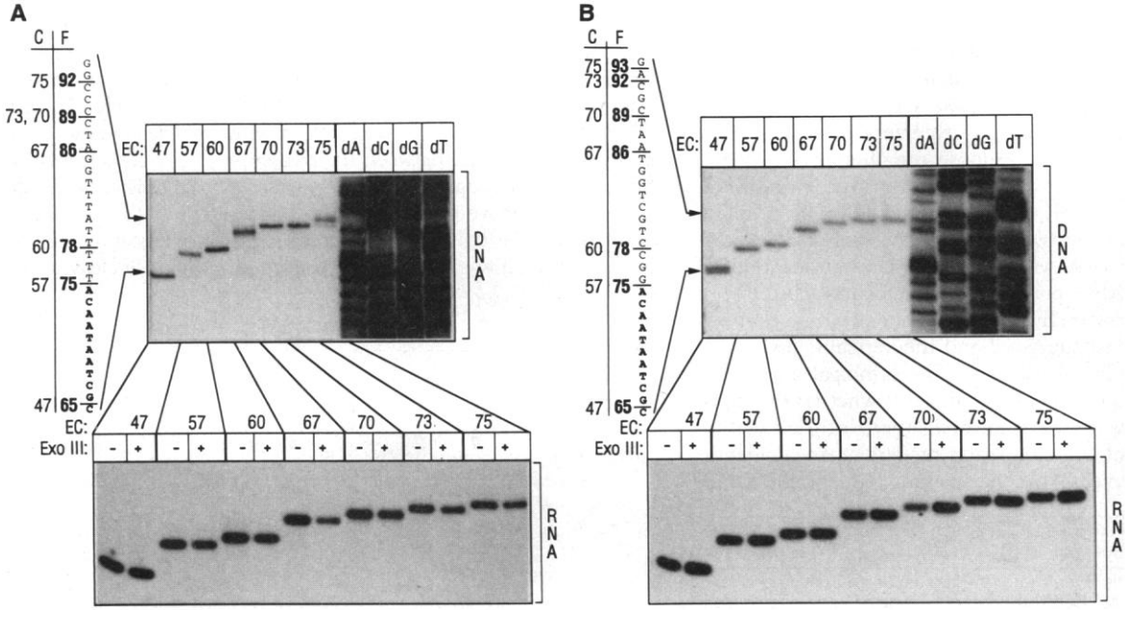

Fig. 4. Effect of sequence change beyond position 75 on the conformation of EC_{73}. RNA polymerase was "walked" through equivalent segments of the two templates. Template 1 (**A**) and template 2 (**B**) are identical in the promoter-proximal sequence (shown in boldface) but differ beyond position 75. The top panels show the protection of DNA from Exo III digestion (*15*). The bottom panels show RNA in each complex analyzed either in the intact form or after treatment with Exo III, as indicated. The tables on the left list the positions of the front end (F) in each complex, which is identified by the length of the transcript (C).

Thus, the switch between the monotonous and the inchworming modes appears to be directed by the nucleotide sequence. The mechanism of this control was examined by comparison of two transcription templates that differed in the DNA sequence beyond position 75. On template 1 (Fig. 4A) an inchworming cycle is seen within the stretch +70 to +75. It is manifested as an invariant position of the front edge in EC_{70} and EC_{73} followed by a 4-nt leap as the complex advances from 73 to 75 (top panel). EC_{73} on template 1 is unstable and yields a nonchaseable dead-end complex, represented by a secondary band in the EC_{75} lane. By contrast, on template 2 (standard template, Fig. 4B) the front edge (top panel) advances concomitantly with the 3′ terminus of RNA (bottom panel). These observations are corroborated by the results of GreB cleavage. As can be seen from Fig. 5 (top panel), starting with EC_{73}, the pattern of transcript cleavage on the two templates is very different. The GreB sensitivity of EC_{73} on template 1 is 10 times higher than on template 2 (bottom panel). Thus, on template 1, EC_{73} displays all the properties of a strained complex, whereas on template 2 it exists in the preferred conformation despite the fact that the transcripts in both cases are identical. These results are summarized in Table 2. We conclude that the signal telling the complex to enter into the inchworming mode is at least partially encoded in the nontranscribed region, ahead of the growing RNA tip.

A model of elongation emerging from these experiments envisages a flexible ternary complex that tends to stay in a preferred, relaxed configuration as it advances along the template. Certain sequence signals encountered on the way cause the complex to rearrange. We propose that the rearrangement is induced by the "anchoring" of the front-end domain to a DNA site in the nontranscribed region accompanied by the cessation of threading of the newly synthesized RNA through the tight RNA-binding site. At the same time, the growth of RNA chain continues, resulting in progressive filling of the loose product-binding site with RNA. The rearrangement is accompanied by buildup of internal strain, which is relieved when the anchoring contacts are broken and the front end leaps ahead with concomitant threading of the transcript through the product site. The length of the leap is sequence specific. We call this parameter "inchworming span." The maximum inchworming span in these experiments is 7, but on some DNA sites it may be as high as 18 (*11*).

Our observations prove the main prediction of the inchworm model that filling and emptying of the loose RNA-binding site and leaping of the front end of the complex are coupled to each other and constitute different manifestations of the same process (*4*, *7*). At the same time, our results reject the key postulate of the original hypothesis that inchworming is intrinsic to elongation and constitutes the very mechanism of movement. Rather we show that inchworming is an incidental event, a response to a signal in the DNA. In the absence of the signal, elongation proceeds monotonously, that is, the polymerase translocates concomitantly with the growth of the transcript. In other words, the inchworming cycle is not the power train converting the

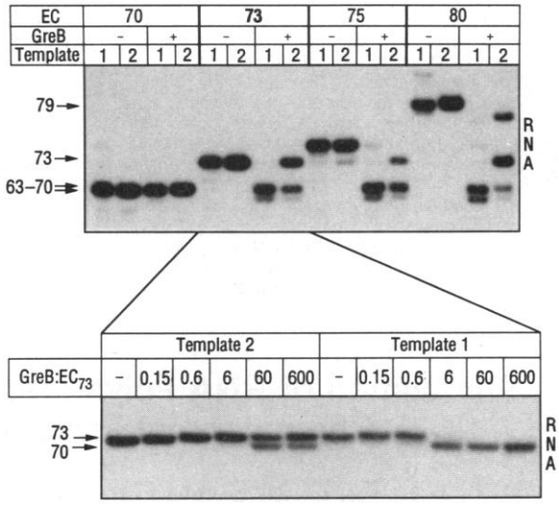

Fig. 5. Effect of sequence change at position 75 on RNA cleavage in elongation complexes. The top panel shows the effect of GreB (180 molar equivalents of GreB per RNA polymerase) on the RNA in the four different complexes. The bottom panel presents the titration of different doses of GreB on EC_{73} on the two templates. The experimental conditions were the same as in Fig. 3 (*16*), except that the Exo III treatment step was omitted.

Table 1. Dimensional parameters of defined elongation complexes. The definitions of the parameters are presented in the text. The table presents the summary of all experimental results obtained for template 2. Strained complexes are shown in boldface. The asterisk marks complexes with a high incidence of dead-end formation under standard conditions and indicates that a mixture of the productive and inactive complex was analyzed in each case. The difference between the four levels of GreB sensitivity represented by ++, +, −/+, − is empirical and spans about three orders of magnitude. It was determined in experiments represented by Figs. 3 and 5 (*17*). The length of the fragment removed by cleavage (R~C) was determined by introducing radioactive phosphate near the 3′ terminus to visualize the cleaved 3′ segment in the gel as described (*4*). The scheme on the right presents the interpretation of the data as alternating phases of movement. ND, not done.

C	F	C~F	GreB sensitivity	R~C	R
20	39	19	-/+	ND	ND
23	**39**	**16**	**++**	**ND**	**ND**
25	**39**	**14**	**++**	**ND**	**ND**
27*	**39**	**12**	**++**	**ND**	**ND**
30	48	18	-/+	1-2	30
32	50	18	-/+	1-2	32
35	53	18	-/+	1-2	35
36	54	18	-/+	1-2	36
40	58	18	-	0	40
44	62	18	-	0	44
47	65	18	-	0	47
50	68	18	-/+	1-2	50
51	69	18	-/+	1-2	51
52	70	18	-/+	1-2	52
53	71	18	-/+	1-2	53
54	**70**	**16**	**++**	**2**	**53**
56*	**68**	**12**	**++**	**4**	**53**
57	75	18	-/+	1-2	57
58	76	18	-/+	1-2	58
59	77	18	-	1-2	59
60	78	18	-	0	60
67	86	19	-/+	1-2	67
70	89	19	-	0	70
73	92	19	-/+	1-2	73
75	93	18	ND	ND	75

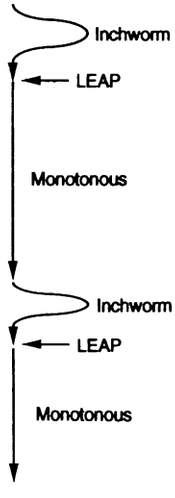

Table 2. Effect of sequence change on dimensional parameters of elongation complexes around position 73. The table summarizes the data of Figs. 4 and 5. The crucial line representing EC_{73} on the two templates is highlighted in boldface. The designations are the same as in Table 1.

	C		F		C~F		GreB sensitivity	
T	1	2	1	2	1	2	1	2
	47	47	65	65	18	18	-	-
	57	57	75	75	18	18	-/+	-/+
	60	60	78	78	18	18	-	-
	67	67	86	86	19	19	-/+	-/+
	70	70	89	89	19	19	-	-
	73	**73**	**89**	**92**	**16**	**19**	**+**	**-/+**
	75	75	92	93	17	18	ND	ND

catalytic reaction into movement but rather is a response of the moving complex to an obstacle encountered on the way (*12*).

The characteristics of the strained elongation complex (shortened C~F distance, filled loose product site, sensitivity to GreB) and its exclusive tendency to convert into the dead-end state find analogy in the phenomenon of elongation arrest by eukaryotic RNA polymerase II (*13*). The arrested Pol II complex is characterized by high sensitivity to the eukaryotic transcript cleavage factor S2, the long register of cleavage, and a shortened distance between the RNA 3′ end and the front edge of the complex. By contrast, elongation-competent Pol II complexes artificially halted elsewhere in the template resemble the preferred conformation of the elongation complex in *E. coli*. These correlations strongly suggest that the model presented here for *E. coli* holds true for the mechanism of transcription elongation in eukaryotic systems.

REFERENCES AND NOTES

1. T. D. Yager and P. H. von Hippel, in Escherichia coli *and* Salmonella Typhimurium. *Cellular and Molecular Biology*, F. C. Neidhardt, Ed. (American Society for Microbiology, Washington, DC, 1987), pp. 1241–1275.
2. B. Krummel and M. J. Chamberlin, *J. Mol. Biol.* **225**, 239 (1992).
3. C. K. Surratt, S. C. Milan, M. J. Chamberlin, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7983 (1991).
4. S. Borukhov *et al.*, *Cell* **72**, 459 (1993).
5. The concept of two binding sites for RNA product has been strengthened by the observation of cleavage of RNA in binary complexes with RNA polymerase [C. R. Altmann, D. E. Solow-Cordero, M. J. Chamberlin, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3784 (1994)].
6. Filling of an internal site with RNA is the concept that was also derived from the observation that the priming nucleotide cross-linked to RNA polymerase can be extended into short RNA chains [A. Mustaev, M. Kashlev, E. Zaychikov, M. Grachev, A. Goldfarb, *J. Biol. Chem.* **268**, 19185 (1993)].
7. M. J. Chamberlin, *Harvey Lectures*, Series 88 (Wiley-Liss, New York, 1994), pp. 1–21.
8. M. Kashlev *et al.*, *Gene* **130**, 9 (1993).
9. K. M. Arndt and M. J. Chamberlin, *J. Mol. Biol.* **213**, 79 (1990).
10. Long-distance backward marching of RNA polymerase after exposure to high doses of cleavage factor S2 have been observed (D. Luse, personal communication). We observed a similar phenomenon after prolonged incubation with high doses of GreB. The reverse movement proceeds through a succession of reverse pause sites that do not coincide with the inchworming signals of forward advancement.
11. E. Nudler, A. Goldfarb, M. Kashlev, unpublished observations.
12. It is worth noting that the distance between the RNA 3′ terminus and the road-blocking psoralene cross-link was reported to be as low as 1 to 2 nt [Y. Shi, H. Gamper, B. V. Houten, J. E. Hearst, *J. Mol. Biol.* **199**, 277 (1988)]. Another relevant situation is that by P. A. Pavco and D. A. Steege [*J. Biol. Chem.* **265**, 9960 (1990)] where the 3′ terminus of the transcript was mapped at the distance of 7 nt before the edge of a road-blocking molecule of cleavage-defective Eco RI endonuclease.
13. W. Gu, W. Powell, J. Mote Jr., D. Reines, *J. Biol. Chem.* **268**, 25604 (1993).
14. The standard template for transcription (template 2) was a polymerase chain reaction (PCR)–amplified 324-bp DNA fragment carrying the T7A1 promoter, in which the first 59 nt of the transcribed region was the same as in template 1 of (*4*). This promoter region was followed by the sequence TG and then by the sequence shown in Fig. 2 and Fig. 4B. His-tagged RNA polymerase was purified and immobilized on Ni-NTA–agarose as described (*8*). In the walking reaction for 12 samples (the experiment of Fig. 2), 20 μl of wet Ni-NTA–agarose pellets were used to immobilize 2 pmol of ribonucleic acid polymerase (RNAP), then 4 pmol of promoter-carrying fragment were added followed by incubation for 5 min at 37°C. The transcription was started by the addition of the priming dinucleotide ApU (0.2 mM), unlabeled adenosine triphosphate (ATP), cytidine 5′-triphosphate, and guanosine triphosphate (25 μM each) and was continued for 10 min at 25°C to make the 20-nt complex (EC_{20}). The reaction was stopped by resin pelleting and washing (*8*), and the RNA in the complex was extended with [α-^{32}P]uridine 5′-triphosphate (3 μl, 300 Ci/mmol, 5 min) to form 3′ end-labeled EC_{21}. Subsequent alternation of the washing and chain extension steps was used for further walking of the complex as described in (*8*). Samples were withdrawn at each walking step in the amount of one-twelfth of the total reaction mixture.
15. The template fragment for Exo III footprinting was obtained by PCR with the use of nonphosphorylated (left) and 5′-phosphorylated (right) 24-nt primers to produce a 5′-OH group in the nontemplate strand for the subsequent enzymatic phosphorylation. The product of PCR was purified from low-melting agarose gel and the nontemplate strand was labeled by T4 polynucleotide kinase and [γ-^{32}P]ATP. The 5′ terminal labeled template was used for the preparation of the stalled elongation complexes as described in (*14*). Exonuclease III (10 U) was added to the immobilized elongation complex in a volume of 10 μl (3 μl of Ni^{2+}-agarose suspension and 7 μl of transcription buffer) and incubated for 6 min at 25°C. The reaction was either stopped by addition of EDTA (50 mM) or Exo III was removed by centrifugation and washing (*14*) and fresh NTPs (250 μM each) were added in "chase" samples (Fig. 2, bottom panel). The Exo III-derived DNA products were separated by 6% denaturing polyacrylamide gel electrophoresis (PAGE).
16. GreB protein was purified as described (*4*). The cleavage reaction of RNA in the immobilized elongation complexes (EC_{52}, EC_{56}, and EC_{70-80}) was carried out in the transcription buffer in 10-μl samples. Each reaction contained ~0.3 pmol of defined ternary complex immobilized on 2 to 5 μl of Ni^{2+}-agarose suspension. GreB was added in the amounts indicated in the figures (180 pmol in the experiment shown in Fig. 5, top panel), and samples were incubated for 10 min at 25°C. For the reaction shown in Fig. 3 (bottom panel), GreB was added to the Exo III–treated complexes after the removal of Exo III protein by centrifugation and washing. All reactions were stopped by addition of EDTA (to 50 mM). RNA products were separated by 6% denaturing PAGE.
17. Because the specific activity of GreB varies substantially from preparation to preparation, the absolute levels of sensitivity to cleavage differ from experiment to experiment (compare Figs. 3 and 5). Thus, the summary data of Table 1 describe a cumulative picture of relative sensitivities of different complexes determined in several experiments.

18. We are grateful to S. Borukhov for providing the GreB protein used in these studies and to other members of the Goldfarb laboratory for helpful discussions. Supported by the NIH grant GM49242 and NSF grant MCB-9218217 (to A.G.). M.K. was supported in part by a long-term fellowship from the Human Frontier Science Program.

11 March 1994; accepted 15 June 1994

Stimulation of RNA Polymerase II Elongation by Chromosomal Protein HMG-14

Han-Fei Ding, Sylvie Rimsky,* Susan C. Batson,† Michael Bustin, Ulla Hansen‡

The high-mobility group protein 14 (HMG-14) is a non-histone chromosomal protein that is preferentially associated with transcriptionally active chromatin. To assess the effect of HMG-14 on transcription by RNA polymerase II, in vivo–assembled chromatin with elevated amounts of HMG-14 was obtained. Here it is shown that HMG-14 enhanced transcription on chromatin templates but not on DNA templates. This protein stimulated the rate of elongation by RNA polymerase II but not the level of initiation of transcription. These findings suggest that the association of HMG-14 with nucleosomes is part of the cellular process involved in the generation of transcriptionally active chromatin.

Transcription in mammalian cells occurs in the context of chromatin, a complex of genomic DNA with histones and non-histone proteins. Consequently, the transcription of a gene can be regulated by alterations in its chromatin structure (*1*). Transcriptionally active chromatin is distinguished from inactive chromatin by alterations in protein composition and by protein and DNA modifications. Candidates for potentiators of the switch from an inactive to an active chromatin structure include proteins associated specifically with active chromatin. The HMG-14 and the closely related HMG-17 proteins are two such candidates, and a variety of experimental data couple these proteins with

H.-F. Ding, S. Rimsky, S. C. Batson, U. Hansen, Division of Molecular Genetics, Dana-Farber Cancer Institute, and Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115, USA.

M. Bustin, Laboratory of Molecular Carcinogenesis, National Cancer Institute–National Institutes of Health, Bethesda, MD 20892, USA.

*Present address: Department of Molecular Biology, Institut Pasteur, 75724 Paris, France.

†Present address: Centre National de la Recherche Scientifique, CHU Purpan, F-31300 Toulouse, France.

‡To whom correspondence should be addressed.