Minimal templates were devised for the efficient generation of yeast RNA polymerase II elongation complexes. A 33-base pair DNA with a 15-residue dC tail at one 3′-end supported the formation of a complex containing the polymerase paused at nucleotide 11 of the duplex region and an RNA of 14–16 residues. The same template could yield an arrested complex with the enzyme at nucleotide 13–15 and RNA of 15–17 residues. These complexes were stable for at least a week under various conditions and could be resolved by gel electrophoresis or purified by ion exchange chromatography. The purified paused complex formed crystals capable of x-ray diffraction to 3.5 Å resolution. The complex remained active in the crystal and, in the presence of nucleoside triphosphates, could efficiently extend the transcript in situ.

RNA chain elongation, once viewed as a monotonous process of ribonucleotide polymerization, now appears mechanistically varied, with important regulatory consequences. RNA polymerase may be unable to extend an RNA chain due to impediments in the DNA template, such as special sequences, damage sites, or protein binding sites (1, 2). The resulting halted complexes may become arrested as a consequence of the polymerase translocating backwards on the template, extruding the 3′-(growing) end of the RNA chain (3, 4). Release from arrest is effected by the elongation factors TFIIH in eukaryotes (5–7) and GreA/B in bacteria (8), which enable the cleavage of extruded RNA and resumption of chain elongation from the newly generated 3′-end (9, 10).

Pausing and arrest afford opportunities for regulatory intervention. Pausing immediately downstream of promoters may be a general rate-limiting step following initiation (11). Gene-specific activator proteins, and also general factors such as elongin, stimulate the elongation process (12–15). Elongin function can play a role in tumorigenesis. Elongin is bound and inhibited by the VHL tumor suppressor protein, and mutations of the VHL protein found in tumors prevent binding (16, 17).

The mechanism and regulation of elongation have been pursued by studies of isolated elongation complexes. In the case of the bacterial enzyme, the complexes are generally formed by promoter-dependent transcription, halted by withholding a single nucleoside triphosphate (18). The halted complexes have been extensively characterized by enzymatic and chemical footprinting, chemical cross-linking, electron microscopy, and other means (19–22). The complexes contain an unwound region ("bubble") of about 15 base pairs of DNA around the nucleotide addition site (20, 23, 24) and nascent RNA hybridized to some extent with the template strand of the DNA (20, 25). Outstanding questions include the paths followed by DNA and RNA, the length of the DNA-RNA hybrid, and the mechanisms of unwinding and rewinding of DNA/DNA and RNA/DNA duplexes.

Answers to these questions will come from determination of the three-dimensional structure of halted complexes. To this end, we have formed halted and arrested complexes with yeast RNA polymerase II. Halted complexes are more difficult to prepare with eukaryotic than bacterial polymerases because initiation at eukaryotic promoters is far less efficient and requires multiple general initiation factors. Here we employ "tailed" templates, bearing 3′-protruding single stranded ends, which support transcription without general initiation factors, starting at a cluster of sites about three residues from the single strand-double strand junction (26). We report on characteristics of polymerase II elongation complexes pertinent to their crystallization and structure determination.

**Experimental Procedures**

**Transcription Templates—**A 220-base pair fragment of pGR220 (kindly provided by Dr. Caroline Kane, University of California, Berkeley), containing a 135-residue T-less cassette, was prepared with a 3′-dC tail by cleavage with AvaI, dephosphorylation with shrimp alkaline phosphatase (U. S. Biochemical Corp.), heat inactivation of the phosphatase, ligation to a phosphorylated synthetic oligonucleotide with an AvaI end and 18 dC residues, cleavage with KpnI, and purification by electrophoresis in low melting agarose, removed by treatment with Gelase (Epicentre Technologies). Plasmids containing shortened T-less cassettes were constructed from the 220-base pair fragment of pGR220 by PCR with the use of oligonucleotide primers containing AvaI and KpnI sites, yielding V51 (from residues 113 to 158 of the pGR220 fragment), V102 (from residues 113 to 206), and V115 (from residues 97 to 206).

Trityl purification of synthetic oligonucleotides (Fig. 1, Anagen Inc.) was sufficient for transcription experiments. Further purification in 6 M urea, 19% polyacrylamide gels was done for purposes of crystallization. Oligonucleotides were electroeluted, dried, dissolved in water, and desalted by filtration through Sephadex G-50. Complementary oligonucleotides, in equimolar amounts on the basis of optical density measurements, were hybridized in 30 mM ammonium sulfate by heating to 80 °C for 5 min and cooling to 20 °C. The presence of a small excess of one strand had no deleterious effect on transcription assays, presumably because their affinity for the polymerase was considerably lower than that of the tailed templates.

**RNA Polymerase II—**Enzyme lacking subunits 4 and 7 was purified from Saccharomyces cerevisiae by immunoadfinity chromatography as described (27) with minor modifications. Polymerase was eluted from the antibody column with 40% propylene glycol instead of glycerol (28). Protein for use in crystallization trials was further purified on either Poros (HE1) heparin or DEAE. Purified polymerase was precipitated by the addition of an equal volume of saturated ammonium sulfate, frozen in liquid nitrogen, and stored at −80 °C.

**Transcription—**RNA polymerase II (0.1–1.5 μg) was incubated with...
a 2- to 3-fold excess of template DNA for 2 min at room temperature in 25 μl of 50 mM Tris, pH 7.5, 6 mM magnesium sulfate, 10 mM DTT, 1 mM mercaptoethanol instead of DTT were incubated with an equal volume of Ni²⁺-nitriloacetic acid resin (Qiagen), previously equilibrated in the reaction buffer, for 10 min at room temperature with gentle agitation. The resin was washed three times (or more if needed to remove radioactivity from the supernatant) by suspension in 1 ml of buffer and centrifugation for 30 s. The resin was suspended in buffer (25 mM Tris, pH 7.5, 40 mM ammonium sulfate, 10 mM DTT, 50 mM zinc sulfate, 50 μM EDTA) by filtration with a Microcon 100. Crystallization was effected as described (31) with the addition of PEG 6000 to lower the ammonium sulfate concentration to 200 mM, preventing dissociation of the complex.

RESULTS

Formation and Stability of Paused Elongation Complexes—Yeast RNA polymerase II lacking subunits Rpb4 and Rpb7 was employed throughout this work because of its homogeneity and its propensity to form both two- and three-dimensional crystals (31). While Rpb4 and Rpb7 are required for promoter-dependent initiation (32), polymerase lacking these subunits is able to transcribe tailed templates and is indistinguishable from wild-type enzyme in RNA chain elongation (not shown). For initial characterization, enzyme with a hexahistidine tag at the C terminus of the largest subunit, Rpb1, was used, but in subsequent work directed toward crystallization, the tag was removed. In previous work directed toward crystallization, the tag was present.

Initial studies employed a plasmid containing a unique AvaI site, adjacent to a 135-bp T-less cassette, terminated by three T residues. Following AvaI digestion and ligation with a synthetic oligonucleotide to introduce a 3'-protruding tail of 18 C residues, transcription in the absence of UTP generated a 135-nucleotide product (Fig. 2), as described by others (25). No transcript from the opposite end of the DNA, also bearing a C tail, was detected, presumably due to the presence of several T residues in the coding strand near this end, resulting in the formation of unstable elongation complexes. The 135-nucleotide RNA resided in stable elongation complexes, as shown by

![Image](image-url)
three further findings. First, the RNA remained associated with the hexahistidine-tagged polymerase following adsorption on Ni\(^{2+}\)-agarose and removal of unbound material. Second, incubation of the Ni\(^{2+}\)-bound elongation complex with UTP extended the transcript by three residues (Fig. 2, first lane), as expected if pausing occurred precisely at the end of the T-less cassette. Third, incubation of the Ni\(^{2+}\)-bound elongation complex with all four nucleoside triphosphates resulted in conversion of most of the RNA to the length expected for runoff at the pause site and the opposite end of the DNA or shorter due to arrest between the pause site and the end (Fig. 2).

The fraction of paused complexes capable of elongation varied from 50 to 95%, depending on the manner of preparation. For example, vigorous vortex mixing or exposure to elevated ionic strength diminished the number of active complexes. Manipulations such as chromatography and filtration, however, had no deleterious effect. Even following washes of Ni\(^{2+}\)-bound halted ternary complex with 2 M potassium acetate, 32% of paused complexes could be elongated (Fig. 2). Complexes were also stable over time, showing no significant loss of activity after a week or more at room temperature (see below).

Minimal Template Size—With a view to eventual crystallization of paused complexes, which might be prevented by protruding DNA, we wished to determine the minimal size of template required for complex formation. DNase I footprint analyses (33), showing protection of 20–24 bp upstream and 19–23 bp downstream of the pause site, suggested the size could be reduced to about 50 bp, if not further. A nested set of templates derived from the T-less cassette, by insertion of restriction sites showed no loss of complex formation, stability, or capacity to elongate upon reduction of the template to a 51-bp fragment with the pause site in the middle, at nucleotide 27 (Fig. 3). Smaller templates were obtained by oligonucleotide synthesis (Figs. 1 and 4). Upon removal of 5 bp from the upstream end and 2 bp from the downstream end of the 51-bp sequence, 47% of the paused complexes were still capable of elongation (Fig. 4). Further removal of 6 and 10 bp from the downstream end had no effect on complex formation but diminished elongation to 9 and 5%, respectively. A length of 22 bp downstream of the halt site was therefore taken as the minimum for the generation of fully active complexes.

Transcripts from shorter templates were better resolved in gels (Figs. 3 and 4), revealing a small degree of length heterogeneity, arising from initiation at a variable location in the single-stranded tail 3–5 residues upstream of the junction with double-stranded DNA. This heterogeneity is a general characteristic of transcription from tailed templates (34). It posed no impediment to crystallization (see below).

Early Arrest Site—All templates derived from the T-less cassette yielded not only the desired products, extending to the halt site, but also a set of shorter RNAs, 15–20 residues in length (Figs. 3 and 4). These shorter RNAs did not derive from abortive transcription and release from the polymerase since they were retained with the hexahistidine-tagged enzyme on Ni\(^{2+}\)-agarose. They evidently resided in arrested complexes since they were not extended by incubation with all four nucleoside triphosphates and since they could be reactivated by the addition of TFII S (Fig. 5).

There were only small changes in the pattern of arrested transcripts arising from the differences in length and sequence of the several templates used. The arrest site was reached when the RNA transcript was 16–18 bases, and taking into
account the variation in transcription start site mentioned above, the polymerase was about 13–15 residues from the beginning of the double-stranded region of the templates. This is in good agreement with the location of early arrest sites reported by others (35). Such early arrest is an intrinsic feature of transcription from tailed templates, which is not found for promoter-dependent transcription by the polymerase and general transcription factors.2 Since heterogeneity in the population of paused complexes formed on tailed templates due to arrest was likely to interfere with crystallization, we sought to avoid the problem by placing the pause site before the early arrest site. The pause site, however, could not be located too far upstream or the transcript would be too short to form a stable complex. These considerations dictated the design of a template (denoted F in Fig. 6) with the pause site at nucleotide 11 of the double-stranded region. Transcription of this template in the absence of UTP yielded an RNA of about 14 residues (Fig. 1, template F, and Fig. 5). Elongation in the presence of all four nucleoside triphosphates was highly efficient, with about 43% arrest at the early site and most of the remainder continuing to run-off. A variant of this template, known to enhance arrest, was prepared with an oligo(T) tract at the arrest site (7, 36) and with additional arrest-inducing sequences upstream of the halt site (template E in Fig. 6) to obtain more uniform behavior during elongation of the paused complex. Transcription again yielded an RNA of about 14–15 residues (Fig. 6), and elongation resulted in almost complete conversion to the arrested complex.

**Fig. 4.** Minimum size of template for paused complex formation. Templates were as in Fig. 1 and as diagrammed on the right, with the number of residues downstream of the first T given. Ni²⁺-agarose was employed as in Fig. 2. Halted (H) complexes and the products of their subsequent elongation (E) are shown. Positions of markers as in Fig. 3 are indicated.

**Fig. 5.** Formation of arrested complexes. Template C was used, and transcription was halted (H) or elongation (E) performed in the presence (+) or absence (−) of TFII S (IIS). Positions of markers as in Fig. 3 are indicated.

**Fig. 6.** Formation of paused and arrested complexes for crystallography. Templates E and F were used, and transcription was halted (H) or transcripts elongated (E). Whereas all complexes halted on template E became arrested during elongation, 45% of complexes formed on template F could transcribe to the end of the DNA.

---

2 D. Price, personal communication.
the ratio of radioactivity to protein was determined. The results for several preparations ranged from 50–95% of the polymerase in paused complexes. This contrasted with previous estimates of 15–55% (18). We therefore sought another approach to the detection and quantitation of paused complexes. We found agarose gel electrophoresis effective for resolving paused complexes from polymerase-template complexes and free polymerase. Paused complexes formed a discrete band, while polymerase-template complexes and free polymerase trailed through the gel, due perhaps to polymerase-agarose interaction (Fig. 7). The gel electrophoretic analysis confirmed the occurrence of a majority of the polymerase in paused complexes.

We turned next to the removal of residual-free (inactive) polymerase. In light of reports that heparin inhibits transcription initiation but not elongation, we thought that heparin-Sepharose might bind free polymerase while allowing elongation complexes to flow through. Indeed, about 10–25% of the polymerase in a paused complex preparation bound to heparin-Sepharose while the remainder, including all the radiolabeled RNA, flowed through (Fig. 8). The RNA to polymerase ratio in the flow-through was 1.13, compared with 0.8 for the material loaded on the column. (As a control, uncomplexed polymerase was applied and virtually all bound to the column.) The efficacy of this procedure was especially apparent when partially proteolyzed polymerase was used. SDS-PAGE revealed the removal of almost all degraded protein from the heparin-Sepharose flow-through (Fig. 9), indicating that only intact polymerase formed stable elongation complexes.

Crystallization of Paused Elongation Complexes—Paused elongation complexes prepared from template E and purified by passage through heparin-Sepharose formed plate-like crystals of approximate dimensions of 150 × 60 × 500 μm, some reaching 800 μm in length. About 10% of the plates were single crystals, and the remainder were multilayered (Fig. 10). Crystals were gradually adapted to mother liquor containing cryoprotectant. Cryoprotected crystals were flash frozen in liquid nitrogen using the flash cooling method (37). Diffraction data obtained with the synchrotron source at CHESS (beamline F1) were complete to about 6 Å and extended in some directions to 3.5 Å resolution. Processing with DENZO revealed a unit cell with parameters a = 197 Å, b = 220 Å, c = 203 Å, and β = 103° and space group C2.

The presence of paused elongation complexes in the crystals was confirmed by incubating a crystal in the presence of ATP, CTP, and [32P]UTP, allowing elongation of existing RNA chains but not the initiation of new chains. The RNA product was detected by gel electrophoresis and autoradiography (Fig. 11). Extensive elongation of RNA in the crystal but not in the mother liquor was observed. Evidently, halted complexes were capable of resuming and completing RNA synthesis in the crystal.

**DISCUSSION**

Novel aspects of this work include the development of tailed templates for forming homogeneous paused and arrested RNA polymerase II transcription complexes, a gel electrophoretic...
method for resolving the complexes, the use of heparin-Sepharose for separation from inactive polymerase molecules, and crystallization of paused complexes in a transcriptionally active state. The high stability of the complexes was anticipated from previous studies of the Escherichia coli enzyme (18). The fraction of active polymerase molecules, however, about 80%, was surprising in light of previous estimates for both yeast and E. coli enzymes of 15–55% (18). The reason for the discrepancy may lie in the methods of estimation. Our figure was obtained from the ratio of RNA to protein in the paused complex preparation, supported by direct visualization in agaro gel slices and by the results of heparin-Sepharose chromatography. The previous numbers relied on quantitation of RNA in gels, which may underestimate the amount, or on measurements of elongation rates, which may vary with the precise conditions used.

Early arrest at a transcript length of about 16 residues is evidently a general property of tailed templates (35). The similarity to the size of the unwound “bubble” in transcription complexes suggests a possible structural basis for arrest. Coding and noncoding strands ordinarily reassociate upstream of the bubble, for example in promoter-driven transcription. Such reassociation can only occur for tailed templates after transcription to about residue 16 and may fail for lack of suitable approximation of the strands, leading to arrest. Subtle differences in the pattern of arrest sites for different templates noted here may be due to the sequence dependence of DNA reassociation and may also reflect template dependence of polymerase conformational changes between elongating and arrested states.

The remarkable outcome of this work is the formation of paused complex crystals capable of diffraction to 3.5 Å resolution. In view of the possibilities for modification of DNA and RNA with heavy atom compounds, as well as structural information forthcoming from studies of the native polymerase and polymerase-DNA complexes, the prospects for structure determination of the paused complex appear favorable. The activity of the paused complex in the crystal, continuing RNA synthesis in situ with no adverse affect on crystal morphology, holds the further promise of structural analysis of the transcription mechanism.

Acknowledgments—We thank Caroline Kane (Univerity of California, Berkeley) for pGR220, Aled Edwards (University of Ontario) for TFIIIS, and Peter David and David Bushnell for help in collecting synchrotron data.

REFERENCES
1. Nudler, E., Kashlev, M., Nikiforov, V., and Goldfarb, A., (1995) Cell 81, 351–357.
2. Landick, R., (1997) Cell 88, 741–744.
3. Gu, W., Powell, W., Monte, J., Jr., and Reines, D. (1993) J. Biol. Chem. 268, 25604–25616.
4. Nudler, E., Goldfarb, A., and Kashlev, M. (1994) Science 265, 793–796.
5. Horikoshi, M., Sekumizu, K., and Natori, S. (1984) J. Biol. Chem. 259, 609–611.
6. Reines, D., Horikoshi, M., and Roeder, R. G., (1987) J. Biol. Chem. 262, 3322–3330.
7. Reines, D., Chamberlin, M. J., and Kane, C. M. (1989) J. Biol. Chem. 264, 10799–10809.
8. Borukhov, S., Sagitov, V., and Goldfarb, A. (1993) Cell 72, 459–466.
9. Izban, M. G., and Luse, D. S. (1990) J. Biol. Chem. 265, 12664–12673.
10. Borukhov, S., Polyakov, A., Nikiforov, V., and Goldfarb, A., (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 899–8902.
11. Krumm, A., Hickey, L. B., and Groudine, M., (1986) Genes Dev. 9, 559–572.
12. Akhtar, A., Faye, G., and Bentley, D. L. (1996) EMBO J. 15, 4624–4634.
13. Yankulov, K., Blau, J., Purton, T., Roberts, S., and Bentley, D. L. (1994) Cell 77, 749–759.
14. Bradsher, J. N., Tan, S., McLaury, H. J., Conaway, J., and Conaway, R. (1993) J. Biol. Chem. 268, 25594–25603.
15. Chafin, D. R., Claussen, T. J., and Price, D. H. (1991) J. Biol. Chem. 266, 9256–9262.
16. Aso, T., Lane, W. S., Conaway, R. C., and Conaway, J. W., Linehan W. M., and Klausner, R. D. (1995) Science 269, 1439–1443.
17. Duan, R. D., Pause, A., Burgess, W. H., Aso, T., Chen, D. Y. T., Garret, K. P., Conaway, R. C., Conaway J. W., Linehan W. M., and Klausner, R. D. (1995) Science 269, 1402–1406.
18. Levin, J. K., Krummell, B. R., and Chamberlin, M. J. (1987) J. Mol. Biol. 196, 85–100.
19. Linn, S. C., and Luse, D. S. (1991) Mol. Cell. Biol. 11, 1508–1522.
20. Lee, D., and Landick, R. (1992) J. Mol. Biol. 228, 759–777.
21. Reese, W. A., Keller, R. W., Vesenka, J. P., Yang, G., and Bustamante, C., (1983) Science 260, 1646–1649.
22. Wang, D., and Landick, R., (1997) J. Biol. Chem. 272, 5989–5994.

3 J. Fu, P. David, D. Bushnell, personal communication.
RNA Polymerase II Elongation Complex Crystals

23. Gamper, H. B., and Hearst, J. E. (1982) Cell 29, 81–90
24. Kainz, M., and Roberts, J. (1992) Science 255, 838–841
25. Rice, G. A., Kane, C. M., and Chamberlin, M. J., (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 4245–4249
26. Kadesch, T. R., and Chamberlin, M. J. (1982) J. Biol. Chem. 257, 5286–5295
27. Edwards, A. M., Darst, S. A., Feaver, W. J., Thompson, N. E., Burges, R. R., and Kornberg, R. D., (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 2122–2126
28. Thompson, N. E., Aronson, D. B., and Burgess, R. R. (1990) J. Biol. Chem. 265, 7069–7077
29. Maniatis T., Fritsch, E. F., and Sambrook, J., (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
30. Bushnell, D. A., Bamdad, C., and Kornberg, R. D. (1996) J. Biol. Chem. 271, 20170–20174
31. Edwards, A. M., Darst, S. A., Hemming, S. A., Li, Y., Kornberg, R. D. (1994) Nat. Struct. Biol. 1, 185–197
32. Edwards, A. E., Kane C. M., Young, R. A., and Kornberg, R. D. (1991) J. Biol. Chem. 266, 71–75
33. Rice, G. A., Chamberlin, M. J., and Kane, C. M. (1993) Nucleic Acids Res. 21, 113–118
34. Dedrick, R. L., and Chamberlin, M. J. (1985) Biochemistry 24, 2245–2253
35. Sluder, A. E., Price, D. H., and Greenleaf, A. L. (1988) J. Biol. Chem. 263, 9917–9925
36. Kerppola, T. M., and Kane, C. M. (1990) Biochemistry 29, 269–278
37. Rodgers, D. W. (1994) Structure 2, 1135–1140
38. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254