AD _____

GRANT NUMBER: DAMD17-97-1-7099

TITLE: Structural Determination of a Transcribing RNA Polymerase II Complex

PRINCIPAL INVESTIGATOR: Averell Gnatt, Ph.D.

RECIPIENT ORGANIZATION: Stanford University
Stanford, California 94305-5401

REPORT DATE: May 1998

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 1998	Annual 5/1/97-4/30/98

4. TITLE AND SUBTITLE
Structural Determination of a transcribing RNA Polymerase II Complex

5. FUNDING NUMBERS
DAMD17-97-1-7099

6. AUTHOR(S)
Averell Gnatt, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Stanford University
Stanford, California 94305-5401

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Material Command
504 Scott Street
ATTN: MCMR-RMI-S
Fort Detrick, MD 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The purpose of the proposed research is to provide a structural basis for understanding the mechanism of transcription, its regulation, and altered regulation as occurs in tumor cells. The goal of the proposed research is to determine the X-ray structure of RNA Polymerase II in the midst of transcribing RNA from a DNA template.

During the first year of funding, template DNA was shortened to a minimal size without affecting the stability of the paused complex or its efficiency of formation. It was determined that previous platelike crystals are sensitive to physical manipulation, soaking in cryoprotectant or soaking with heavy metal clusters. Crystallization trials of polymerase resulted in new non-platelike crystal forms allowing for the collection of a preliminary data set to 7.5 Angstroms. This is the best elongation complex data set, since previous platelike crystals had mosaic spreads over 1 degree. Recently new crystallization conditions allowed the growing of crystals in cryoprotectant and the generation of yet new crystal forms. Growing in cryoprotectant is a major accomplishment considering the sensitivity of the crystals to manipulations.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Regulation, RNA Polymerase II, crystallography, elongation, ternary complex.	14
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

✓ Where copyrighted material is quoted, permission has been obtained to use such material.

✓ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

✓ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

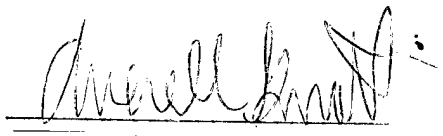
✓ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.



PI - Signature _____ Date

TABLE OF CONTENTS

Structural Determination of a Transcribing RNA Polymerase II complex.

Introduction

The goal of the proposed research is to determine the X-ray structures of RNA polymerase II elongation complex at atomic resolution and with regulatory proteins. The purpose of the proposed research is to provide a structural basis for understanding the mechanism of transcription, regulation of the process, and altered regulation as occurs in tumor cells. The problem is challenging, since the polymerase alone comprises 15 polypeptides with a total mass of nearly 600,000 Dalton, and addition of the auxiliary factors adds a further level of complexity.

It has been found that mutations in a tumor suppressor gene, P^{53}, is directly associated with breast cancer. Currently, evidence indicates that P^{53} is a transcriptional regulatory protein. Treatment of breast cancer involves radiation therapy, chemotherapy and surgery. In devising an effective treatment, it would be helpful to understand the underlying cause/s of tumorigenesis. Since tumors are cells whose regulation has been altered, it is imperative to uncover the normal regulatory mechanism and those that occur in cancer cells. As such, this proposal can create opportunities for improved diagnosis, prevention, and treatment.

Although mutations in genes can predispose individuals for breast cancer, the exact mechanism that leads cells to obtain tumorous properties is not fully understood. The path by which mutations are capable of altering cellular traits is by altering the regulation of specific genes. The point of altered regulation can be at initiation or elongation of RNA polymerase II. During elongation, RNA polymerase II pauses on its transcript. Proteins such as TFIIS, can regulate the amount of the read through. Indeed, an additional elongation factor, SIII, has been shown to be a target of the VHL tumor suppressor protein and able to directly regulate its function (1). Mutations in the VHL gene predispose individuals to a variety of tumors (2). This is a clear case of point mutations in a gene, directly affecting the regulatory mechanism.

Eukaryotic cells contain three forms of RNA polymerase, designated I, II, and III (or A, B, and C), made up of 10-15 polypeptides. Despite the many subunits, the eukaryotic enzymes are alone incapable of recognizing a promoter and initiating transcription; a number of accessory factors, referred to as general transcription or general initiation factors, are required (for reviews see ref. 3, 4, 5). The diversity of RNA polymerases belies a remarkable degree of conservation across species. The two largest subunits, which together account for two-thirds of the mass of the polymerases, are homologous in sequence from *E. coli* to man. General transcription factors

and activator proteins are also conserved and even interchangeable between yeast and human cells. The transcription mechanism appears to be universal: the mechanistic details of initiation, RNA chain elongation, and termination, insofar as they have been revealed, are fundamentally the same in bacteria, yeast and higher eukaryotic cells. Recently, 6 of the 7 human RNAPII subunits tested were found able to replace their yeast counterparts in vivo (6). Therefore, studies of yeast RNAPII may be expected to reveal general principles of eukaryotic transcription and its regulation. The yeast enzyme is especially suited for 3-D structural analysis because a large amount of pure material can readily be obtained from yeast cell culture. No other cell culture system is able to fulfill such a requirement. Yeast RNA polymerase lacking subunits subunits 4 and 7 were originally used for 2-dimentional crystallography since they were more homogenous than the wild type enzyme (7,8).

Although, RNA polymerase II requires a set of additional protein factors to start transcription at a promoter, the polymerase with or without subunits 4 and 7 is capable of initiation on a single strand protruding from the 3'-end of duplex DNA (Kade9). Such a "tailed" template may be viewed as half of an unwound "bubble", known to occur at the active site of RNA polymerase molecules during transcription. Consistent with this idea, transcription starts within the single stranded region, about three bases from the junction with duplex DNA, in both tailed templates and in the unwound bubble of an elongation complex (10). At least two possible paused complexes can exist. The first is halted due to the lack of a single nucleotide such as UTP (11). The second is arrested even in the presence of all nucleotides, due to the DNA structure arising from its primary sequence (12). Although structural determination of the ternary complex in a functional state until date has not been shown, use of tailed templates has allowed the applicant to develop a system with appropriate templates for the generation, purification and crystallization of this complex (13).

RNA polymerase in the act of transcription is essentially irreversibly bound to both the DNA template and the nascent RNA chain (no dissociation occurs even upon treatment with a detergent such as sarkosyl). Such "elongation" complexes formed with *E. coli* RNA polymerase (14) remain fully active in transcription following exposure to concentrated salt solution (1M NH_4Cl) or storage for long periods (5 days at 4^oC). These characteristics are conducive to crystallization, and prompted the original crystallization of elongation complexes prior to this fellowship. Elongation complexes were "halted" on tailed templates by transcription in the absence of UTP, so the polymerase halted when the first T residue in the template was reached. The halted complexes generated on the tailed templates are advantageous for crystallization because of their uniformity in content of DNA and RNA sequences (13).

Before acquiring this fellowship, platelike crystals of the delta 4/7 enzyme were successfully grown and diffracted (13). Diffraction data taken on station F-1 at CHESS were complete to 6.4 Å and extended in some directions to 3.5 Å (Fig. 3). Few crystals survived soaking and freezing, however, and the mosaic spread of the diffraction was high (greater than 1°).

Body

Several tasks were set forth in the original proposal to be performed during the first year the duration of this fellowship. These are listed below.

Task1. 1-12 months. Continued crystallization of ternary complex, which currently yield platelike crystal. This will be done using in house X-ray diffraction equipment as well as synchrotron radiation from the Stanford Linear Accelerator (SLAC).

Task2. 1-18 months. Derivitization of current crystals with site specific heavy metal clusters for initial phases. This is to be performed employing undecagold heavy metal clusters currently available, and novel clusters as they become available. Tantalum clusters have just been employed on non-ternary polymerase and look quite promising.

Task 3. 9-12 months. Improving current crystals by trying to shorten present DNA templates, and crystallization of complexes with templates of different sequences. Several new template sequences are currently in hand to use. Shortening these by several bases should suffice to improve crystal form. Each novel template will be employed in generating ternary complex. It will then need to be studied biochemically in a system that the applicant has developed. This will ensure that the new complexes are stable and functional.

Figure 1. Templates used in generating elongation complexes

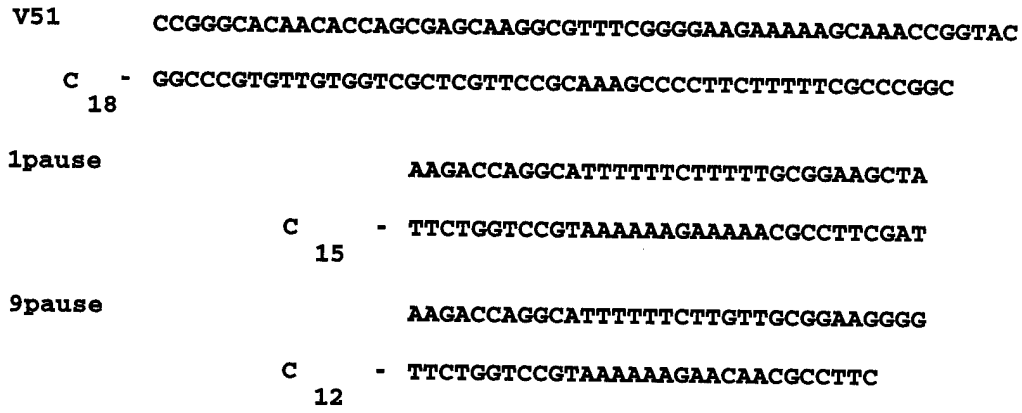


Figure 1 Tailed template DNA's for formation of arrested RNA polymerase II elongation complexes. Templates are aligned at their halt sites which is the first T in the poly -T tract

Templates Employed in Generating Elongation complex.

In the original proposal a 51 base pair template, V51 (fig. 1) and was shown to be efficient for the generation of homogeneously paused elongation complexes using yeast RNA Polymerase II. Withholding UTP at a poly T site on the template was the key in generating these complexes. The exact template employed for generating ternary complexes for crystallization, template 1pause (figure 1) was based on the same overall structure of V51. Recently with task 3 (see above) in mind this template strand was further shortened by 6 bases (termed 9pause) with no effect on efficiency of transcription or stability of the elongation complex. These changes included shortening of the poly dC tail from 15 to 12 bases and reducing the 5′ terminal of the coding strand by 3 bases allowing for a 3 base overhang by the noncoding strand (fig 2). This was performed to study the affect of DNA size on resolution of diffracted crystals. An additional change which had no apparent affect was the alteration of one of the base 21 (relative to the first double stranded nucleotide transcribed) from T to G (fig. 1). Nuclease protection assays of a 150 base pair ternary complex shows that approximately 45 bases of DNA is protected by the polymerase (11). Therefore further shortening of the downstream end of the shortened DNA template is within the protected region. Decreasing the single stranded tail further than 12 bases is not feasible since that would drastically decreases initiation (9).

The halted elongation complex generated by withholding UTP could be efficiently extended (fig. 2) after adding UTP, as efficiently as was done for template 1pause (13). It also allowed for a nearly complete arrested complex which was unable to elongate past the poly T tract in the template.

Figure 2. RNA in elongation complexes employing template 9pause

Platelike crystals of RNA polymerase II elongation complex.

In the original proposal, X- ray diffraction studies allowed for a complete data set of an elongation complex to 7 Å (Table 3). This was done using yeast RNA polymerase II which was deficient in two subunits four and seven (delta 4/7). In implementing the first and second tasks (above) which was to use the "platelike" elongation crystals to achieve phasing information, several difficulties were encountered.

Figure 2 Formation of paused (denoted "H" for halted) and arrested (denoted "E" for elongated) RNA polymerase II elongation complexes on the template in Fig. 1. Transcripts were analyzed by gel electrophoresis and autoradiography. Multiple bands were obtained due to the well known lack of precision of initiation on tailed templates.

1. Sensitivity to manipulation. It was necessary to freeze the polymerase crystals since diffraction from capillary mounted crystals decay within 2 hours. A solution of mother liquor and propylene glycol (16% final concentration) was the least damaging of the cryoprotectants studied. Even then most crystals cracked, dissociated or displayed a decreased ability to diffract resulting in only low resolution diffraction (6-7 Å) with a high mosaic spread of greater than 1 degree. For example the complete data set mentioned in the original proposal was 71% complete to 6.4 Angstroms but had a mosaic spread of 1.5 degree.

2. Sensitivity to heavy metal complexes. To achieve phasing data for a protein the size of polymerase, heavy metal clusters were employed. Recently, co- workers successfully employed a Tantalum (Ta6Br122+) complex in the Kornberg lab, in achieving phase information for an RNA Polymerase II not in ternary complex. When soaking the Delta 4/7 elongation crystals with Ta6Br122+ most crystals cracked. This behavior is clearly different than the

non-ternary enzyme and strongly suggests that the Ta6Br122+ complex binds to a region that is occupied by nucleic acids or that it interacts with nucleic acids in the elongation complex. In either case a disruption of the crystal lattice and cracking of the crystals prior to soaking in cryoprotectant was observed.

3. Although 1/15 of the platelike crystals do diffract to high resolution, the inability to manipulate them and their high mosaic spread suggests that either better freezing conditions are necessary or a different crystal form is needed. Therefore, a variety of crystallization conditions were tested in addition to conditions that allowed for high resolution diffraction of the polymerase alone. The non- ternary polymerase crystals had an I222 symmetry whereas the ternary plate crystals are C222. These attempts result in only platelike crystals. It appears that the additional DNA and RNA in the complex prevent specific protein- protein interactions necessary for the generation of the I222 crystal forms. The search for conditions was somewhat limited though, since the use of high salt causes the paused elongation complex to be unable to further extend the RNA chain (13). For that reason PEG 6000 has been employed as the main precipitating agent as opposed to having high salt concentrations.

Some valuable information was gained though by the addition of additives and/or screening of crystallization conditions, and are summarized below. Despite this no improvement in the quality of the platelike crystals was observed.

Figure 2. Additives to Mother Liquor that affect crystal growth

Additive	Effect
Spermine	Slightly decreases ternary complex solubility
Sodium Chloride	Strongly decreases ternary complex solubility
Ethanol	Slightly decreases ternary complex solubility
Cadmium Chloride	Increases the size of crystals, may substitute for Zinc in Zinc fingers or bind Nucleic acids in the elongation complex stabilizing the complex.
Lutetium Sulfate	Increased the size of crystals though not as much as Cadmium Chloride did.

Wild type RNA polymerase II elongation complex.

As described above, a search for a crystal form that could be easily manipulated was called for. A co- worker has recently crystallized the wild type enzyme containing subunits 4 and 7 and has shown diffraction before freezing (capillary mounted crystals) to high resolution and after freezing to low resolution (6.0 Å). These crystals were orthorhombic C222 with different

cell unit dimensions than the elongation plate crystals (fig 3). Ternary complexes of wild type polymerase were then generated, and appear to be biochemically identical (not shown) to those of the delta 4/7 enzyme. After initial crystallization non-platelike crystals with four different apparent morphologies were successfully grown. Initial diffraction at SSRL beam line 9-1 of several of these crystals shows them to be similar in morphology, cell unit and symmetry to wild type crystals.

As a result many crystals were frozen, and some conditions which contained PEG 6000 and ammonium phosphate as buffer proved to allow for diffraction. These crystals are not as sensitive to soaking with cryoprotectant as the plate crystals were. Diffraction in these crystals though has been observed to be limited to 6 Å even after screening of 180 crystals with an in house rotating anode X- ray source. In addition all four different crystal morphologies appear to have similar cell unit dimensions. A full data set of one of one of these crystals has recently been accomplished to 7.5 Å resolution. Computational analysis of the data was performed with MOSSFLM (integration) and scaling with SCALA (from the ccp4 package). The mosaic spread of this crystal (0.75 degrees) is far lower than that of the delta 4/7 crystal data set (1.5 degrees). This improvement may be crucial in allowing for a low resolution structure by molecular replacement using the non-ternary delta 4/7 enzyme that has been successfully phased to 6.0 Å.

It appears that freezing conditions may play a key role with the wild type elongation complexes as well. Although after soaking in propylene glycol cracking of crystals was not observed, a loss of birefringence of most of the crystals occurred. This suggests that better freezing conditions may be necessary for both the wild type and the delta 4/7 enzyme. Recently, screens for such conditions look quite promising, as detailed below.

Improving crystallization and freezing conditions

Since elongation complexes may dissociate with high salt, the ideal crystallization conditions found to date included precipitating agents such as polyethelyneglycol 6000 and ammonium phosphate as salt up to 300 mM. It would though, be ideal if crystals could be grown in cryoprotectant. This would severely minimize the damage, which appears to be a major constraint in achieving high resolution diffraction.

A screen was set up employing wild type elongation complex, PEG400 as precipitant and ammonium phosphate (up to 300 mM) as salt/buffer. The logic behind employing PEG400 is that above 20% it is a good cryoprotectant, and elongation complex crystallized well in PEG6000. Elongation complexes in these trials remained in solution for months (until this day) and did not precipitate out or form crystals. This clearly indicated that PEG400 is not an efficient precipitating agent even at 25%.

Figure 3. Crystallographic data for yeast RNA polymerase II (pol II) and its complexes

Molecule	Unit Cell	Space Group	X-ray Source	Resolution Limit (Å)	R_{symm}	< I/sigma >	Completeness
pol II	a = 131.3 b = 223.7 c = 368.9	I222	SSRL BL7-1	3.0	9.9 % for 50.0-4.0 Å	30.0	75% for 50-4.0 Å
	a = 211.4 b = 222.3 c = 321.6	I222 or $I2_12_12_1$	SSRL BL7-1	5.0	7.5% for 50.0-5.0 Å		98% for 50.0-5.0 Å
Transcription elongation complex	a = 174.9 b = 222.2 c = 196.5 β = 103.3°	C2	CHESS F-1	6.0	6.1% for 70-6.4 Å		71% for 70-6.4
wild-type pol II	a = 218.9 b = 390.5 c = 279.8	C222	SSRL BL 1-5	6.0	6.3% for 50-6.0 Å	4.7	95% to 6.0 Å
*Wild-type elongation complex	a = 219.8 b = 389.1 c = 272.0	C222	SSRL BL 1-5	6.0	7.6% for 23.7-7.5 Å	7.2	74% to 7.5 Å

* Performed under the current study.

After these observations, and the considerable amount of screening of elongation complex crystallization conditions, several elements were then combined in an attempt to increase precipitation of the ternary complex. Both sodium chloride and PEG 6000 were shown to be good precipitating agents in past screens. Several additional screens were then set up all employing 20% PEG 400, 100mM ammonium phosphate pH 6.0 as buffer and either PEG6000 or sodium chloride as precipitant. Small crystals were observed in many of the resultant trials. They all appear to have a single elongated morphology, different than those grown in the absence of PEG 400. The largest of those crystals were still relatively small, 100 X 50 X 40. Considering the large size of polymerase and the small size of the crystal diffraction was not expected to be observed. Some of these crystals were directly and successfully frozen. This is the first time crystals of elongation complex can be grown in cryoprotectant. Although only a few of these crystals were diffracted and despite their size weak diffraction to 10 Angstroms was observed on an in house rotating anode equipped with a Siemans area detector.

These crystals though appear to hold promise in that they are grown in cryoprotectant and can be directly frozen. At the time of the submission of

this report, several new crystal trials have been set up. As mentioned above adding cadmium chloride to elongation crystals increased crystal size dramatically. This may allow for the elongation complex crystals grown in PEG 400 to achieve the size necessary for good diffraction.

Conclusions

The tasks set out for the first year included employing platelike crystals of delta 4/7 enzyme and heavy metal clusters for phasing. In addition improvement of the current template by shortening was called for. In the second and third years the goal will be an attempt at high resolution structural studies. Indeed the current study was successful in attaining the first year goals and more.

1. It was determined that the platelike crystals are sensitive to manipulations including soaking in cryoprotectant and heavy metal clusters, and that the nucleic acids in the complex are most likely preventing protein- protein interactions necessary for growth of the I222 crystal forms of non- ternary polymerase which diffract to high resolution.

2. Approximately 45 bases are well protected by RNA polymerase II (11). The DNA template originally employed for elongation complex was further shortened without any deleterious affect. Shortening beyond this limit may not allow for efficient elongation after pausing (13). This could always raise the question of the biochemical integrity of the complex. At this point the current shorter template will be employed as the template of choice for all further studies.

3. A data set was achieved to low resolution (7.5) of the wild type elongation complex. The mosaic spread is approximately 0.75 degrees. The original delta 4/7 platelike crystals had a mosaic spread of 1.5 and the new crystals contain protein with subunits 4 and 7 that are required in specific initiation of transcription, though not necessary for elongation. The current data is of the highest quality collected to date on elongation complexes of RNA Polymerase II.

4. Phase information to low resolution from the non-ternary RNA polymerase II is now at hand (Jianhua Fu, personal communication). This data may prove useful in molecular replacement with the wild type elongation complex to arrive at a low resolution map. This would have been difficult using the delta 4/7 crystal data collected prior to this study, considering its high mosaic spread. As a result of the research done under the first year of this fellowship X- ray structural analysis of an elongation complex of RNA polymerase II may now be attempted to low resolution.

5. New crystal growth conditions have now been found that include growth in cryoprotectant. This may allow for the circumvention of the main

problems involved in crystal manipulation that has hampered both the quality of data derived from both the delta 4/7 and the wild type elongation crystal. It will be necessary to further capitalize and improve these conditions for both delta 4/7 and native elongation complexes, to allow for larger crystals. This effort is now underway.

References

1. Aso, T., Lane, W.S., Weliky, Conaway, J.W., & Conaway, R.C. **Science 269:**1439-1443 (1995).

2. 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., & Klausner, R.D. **Science 269**:1402-1406 (1995).

3. Sawadogo, M., & Sentenac, A. **Ann. Rev. Biochem. 59**:711-754 (1990).

4. Conaway,. J.W., & Conaway, R.C. **J. Biol. Chem. 266**:17721-17724 (1991).

5. Young, R.A. **Ann. Rev. Biochem. 60**:689-715 (1991).

6. Mckune, K. Moore, P.A. Hull, M.W. & Woychik, N. A. **Molecular Cell Biology 15:**6895-6900 (1995).

7. Edwards, A.E., Kane C.M., Young, R.A. & Kornberg, R.D. **J. Biol. Chem. 266**:71-75 (1991).

8. Edwards, A.M., Darst, S.A., Feaver, W.J., Thompson, N.E., Burges, R.R., & Kornberg, R.D., . **Proc. Natl. Acad. Sci. USA 87:**2122-2126 (1990).

9. Kadesch, T.R. & Chamberlin, M.J. . **J. Biol. Chem. 257:**5286-5295 (1982).

10. Dedrick, R.L., Chamberlin, M. J. Studies on transcription of 3'-extended templates by mammalian RNA polymerase II. Parameters that affect the initiation and elongation reactions. **Biochemistry 24**:2245-2253 (1985).

11. Rice, G.A., Chamberlin, M. J., & Kane, C.M. **Nucleic Acids. Res. 21**:113-118 (1993).

12. Nudler, E., Kashlev, M., Nikiforov, V., & Goldfarb, A., **Cell 81**:351-357 (1995).

13. Gnatt, A., Fu, J., & Kornberg, R.D. **J. Biol. Chem.** 272:30799-30805 (1997).

14. Levin, J.R., Krummell, B.R., & Chamberlin, M.J. . **J. Mol. Biol. 196**:85-100 (1987).