Article

T7 RNA Polymerase Transcription with 5-Position Modified UTP Derivatives

Jonathan D. Vaught, Torin Dewey, and Bruce E. Eaton

More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S

A R T I C L E S

T7 RNA Polymerase Transcription with 5-Position Modified UTP Derivatives

Jonathan D. Vaught, Torin Dewey, and Bruce E. Eaton*

Contribution from the College of Physical and Mathematical Sciences, Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695

Received February 21, 2004; E-mail: Bruce_Eaton@ncsu.edu

Abstract: Seven UTP derivatives modified at the 5-position through an amide linkage were tested as substrates for T7 RNA polymerase (T7 RNAP) transcription. All UTP derivatives gave good yields of full-length transcript even from DNA templates that showed a significant number of abortive transcripts using unmodified UTP. A kinetic assay to determine the relative K_m and V_{max} for T7 RNAP transcription gave surprisingly similar values for UTP and the 5-position hydrophobic modifications phenyl, 4-pyridyl, 2-pyridyl, indolyl, and isobutyl. The 5-position modifications imidazole and amino, which could both be positively charged, gave K_m values significantly higher than UTP. All seven UTP derivatives gave relative V_{max} values similar to UTP, indicating that insertion of these modified bases into the transcript did not impede its elongation.

Introduction

Recently, in vitro selection[1,3] has been used to expand the scope of RNA catalysis beyond nucleic acid phosphodiester bond chemistry. Numerous RNA and DNA catalysts for organic reactions have now been discovered through in vitro selection techniques. Examples to date include acyl transfers,[4-9] amide synthesis,[10] Diels–Alder cycloadditions,[11-15] Michael additions,[16] nucleophilic substitutions,[17] porphyin metalations,[18,19] enzyme cofactor synthesis,[20,21] and urea synthesis.[22] There appears to be an ever expanding list of reactions that RNA can catalyze. However, despite these apparent successes in finding

new RNA catalysts for organic reactions, their efficiency is relatively low when compared to protein enzymes, and their relevance to either biochemistry or synthetic chemistry is still in question.

A serious limitation of RNA and DNA enzymes compared to proteins is that they lack the diversity of functional groups typically found in active sites. It might be possible to increase the catalytic potential of RNA and DNA if added functionality could be included. Recent advances in the chemical and enzymatic modification of RNA and DNA are being developed to increase the structural and functional diversity in nucleic acids.[2,23-34] Using RNA libraries for in vitro selection experiments is an efficient way to determine if RNA modification will yield new or improved catalysts. However, to be useful in

(1) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505−510.
(2) Dewey, T. M.; Mundt, A. A.; Crouch, G. J.; Zyzniewski, M. C.; Eaton, B. E. *J. Am. Chem. Soc.* **1995**, *117*, 8474−8475.
(3) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818−822.
(4) Lohse, P. A.; Szostak, J. W. *Nature* **1996**, *381*, 442−444.
(5) Suga, H.; Lohse, P. A.; Szostak, J. W. *J. Am. Chem. Soc.* **1998**, *120*, 1151−1156.
(6) Lee, N.; Bessho, Y.; Wei, K.; Szostak, J. W.; Suga, H. *Nat. Struct. Biol.* **2000**, *7*, 28−33.
(7) Illangasekare, M.; Sanchez, G.; Nickles, T.; Yarus, M. *Science* **1995**, *267*, 643−647.
(8) Illangasekare, M.; Yarus, M. *J. Mol. Biol.* **1997**, *268*, 631−639.
(9) Kumar, R. K.; Yarus, M. *Biochemistry* **2001**, *40*, 6998−7004.
(10) Wiegand, T. W.; Janssen, R. C.; Eaton, B. E. *Chem. Biol.* **1997**, *4*, 675−683.
(11) Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E. *Nature* **1997**, *389*, 54−57.
(12) Tarasow, T. M.; Tarasow, S. L.; Tu, C.; Kellogg, E.; Eaton, B. E. *J. Am. Chem. Soc.* **1999**, *121*, 3614−3617.
(13) Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E. *J. Am. Chem. Soc.* **2000**, *122*, 1015−1021.
(14) Seelig, B.; Jaschke, A. *Chem. Biol.* **1999**, *6*, 167−176.
(15) Stuhlmann, F.; Jaschke, A. *J. Am. Chem. Soc.* **2002**, *124*, 3238−3244.
(16) Sengle, G.; Eisenfuhr, A.; Arora, P. S.; Nowick, J. S.; Famulok, M. *Chem. Biol.* **2001**, *8*, 459−473.
(17) Wecker, M.; Smith, D.; Gold, L. *Rna-A Publication of the Rna Society* **1996**, *2*, 982−994.
(18) Li, Y. F.; Sen, D. *Nat. Struct. Biol.* **1996**, *3*, 743−747.
(19) Li, Y. F.; Sen, D. *Biochemistry* **1997**, *36*, 5589−5599.
(20) Huang, F. Q.; Bugg, C. W.; Yarus, M. *Biochemistry* **2000**, *39*, 15548−15555.
(21) Jadhav, V. R.; Yarus, M. *Biochemistry* **2002**, *41*, 723−729.
(22) Nieuwlandt, D.; West, M.; Cheng, X. Q.; Kirshenheuter, G.; Eaton, B. E. *ChemBioChem* **2003**, *4*, 651−654.

(23) Benner, S. A.; Battersby, T. R.; Eschgfaller, B.; Hutter, D.; Kodra, J. T.; Lutz, S.; Arslan, T.; Baschlin, D. K.; Blattler, M.; Egli, M.; Hammer, C.; Held, H. A.; Horlacher, J.; Huang, Z.; Hyrup, B.; Jenny, T. F.; Jurczyk, S. C.; Konig, M.; von Krosigk, U.; Lutz, M. J.; MacPherson, L. J.; Moroney, S. E.; Muller, E.; Nambiar, K. P.; Piccirilli, J. A.; Switzer, C. Y.; Vogel, J. J.; Richert, C.; Roughton, A. L.; Schmidt, J.; Schneider, K. C.; Stackhouse, J. *Pure Appl. Chem.* **1998**, *70*, 263−266.
(24) Dewey, T. M.; Zyzniewski, M. C.; Eaton, B. E. *Nucleosides Nucleotides* **1996**, *15*, 1611−1617.
(25) Gourlain, T.; Sidorov, A.; Mignet, N.; Thorpe, S. J.; Lee, S. E.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001**, *29*, 1898−1905.
(26) Lutz, S.; Burgstaller, P.; Benner, S. A. *Nucleic Acids Res.* **1999**, *27*, 2792−2798.
(27) Grahn, E.; Stonehouse, N. J.; Adams, C. J.; Fridborg, K.; Beigelman, L.; Matulic-Adamic, J.; Warriner, S. L.; Stockley, P. G.; Liljas, L. *Nucleic Acids Res.* **2000**, *28*, 4611−4616.
(28) Perrin, D. M.; Garestier, T.; Helene, C. *Nucleosides Nucleotides* **1999**, *18*, 377−391.
(29) Perrin, D. M.; Garestier, T.; Helene, C. *J. Am. Chem. Soc.* **2001**, *123*, 1556−1563.
(30) Sakthivel, K.; Barbas, C. F. *Angew. Chem., Int. Ed.* **1998**, *37*, 2872−2875.
(31) Teramoto, N.; Imanishi, Y.; Ito, Y. *Bioconjugate Chem.* **2000**, *11*, 744−748.
(32) Thum, O.; Jager, S.; Famulok, M. *Angew. Chem., Int. Ed.* **2001**, *40*, 3990−3993.
(33) Vaish, N. K.; Fraley, A. W.; Szostak, J. W.; McLaughlin, L. W. *Nucleic Acids Res.* **2000**, *28*, 3316−3322.
(34) Lermer, L.; Roupioz, Y.; Ting, R.; Perrin, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 9960−9961.

Scheme 1. Steps of the RNA Catalysts Selection Cycle
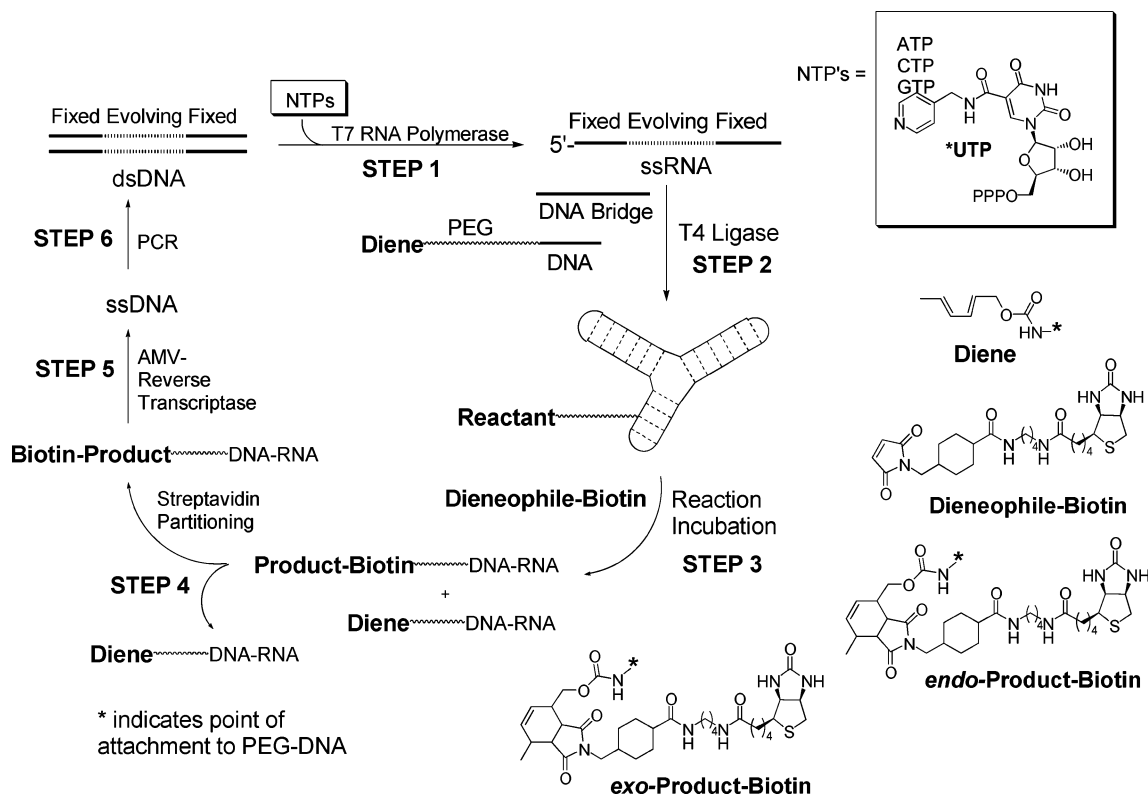
an in vitro selection the modification of the nucleic acid must be accepted by the enzymes used in the transcription steps of the selection process. Herein we focus on 5-position modified UTP derivatives and their incorporation into RNA by T7 RNA polymerase (T7 RNAP) transcription.

RNA Catalyst Selection. In vitro selection for RNA catalysts begins with a library of 10^{14} sequences, generated by chemical synthesis at random, which are then subjected to repeated cycles of selection where active RNA catalysts become enriched in the population. The selection cycles require the use of enzymes for transcription and reverse transcription. The in vitro selection cycle used previously for discovering the first RNA Diels–Alderases is shown in Scheme 1. The process begins with ssDNA library of 10^{14} sequences made on an automated DNA synthesizer. The ends of the ssDNA library contain fixed regions required for the enzymes used in the amplification, and the center of the ssDNA contains a region of random sequence, typically 100 nucleotides in length. To prepare the RNA library, a starting dsDNA template for T7 RNAP transcription is made by two-cycle PCR of the chemically synthesized ssDNA library.

The key step relevant to the results reported herein is Step 1 of Scheme 1 where a modified RNA library is created by T7 RNAP transcription of a dsDNA template using ATP, CTP, GTP, and derivatives of UTP. One barrier to expanding this list of modified RNA libraries stems from the fact that if there is a substantial bias during T7 transcription against incorporation of modified bases into the RNA library or significant mis-incorporation, the potential sequence space (10^{14}) possible and concomitant active sites will be significantly underrepresented. Therefore it is important to understand whether modified nucleotides are suitable substrates for T7 RNAP. It had been previously reported that 5-benzoyl-UTP (5-BU) was a suitable substrate for T7 RNAP;[2] however, no successful selections using

this modification have been reported. In addition 5-position pyridyl **2** and imidazoyl **4** UTP derivatives have been used successfully for in vitro selection experiments to give new catalysts,[10,11,22] but the details of how they function as substrates for T7 RNAP has not been reported. Herein we examine an expanded list of UTP derivatives, including **2** and **4** all connected to the 5-position of uridine through an amide linkage (Figure 1).

The synthesis of UTP derivatives **1**–**7** was accomplished by a previously reported Pd catalyzed carboxyamidation procedure[24] followed by the Eckstein triphosphate synthesis.[35] The carboxyamidation method is attractive for creating UTP modi-
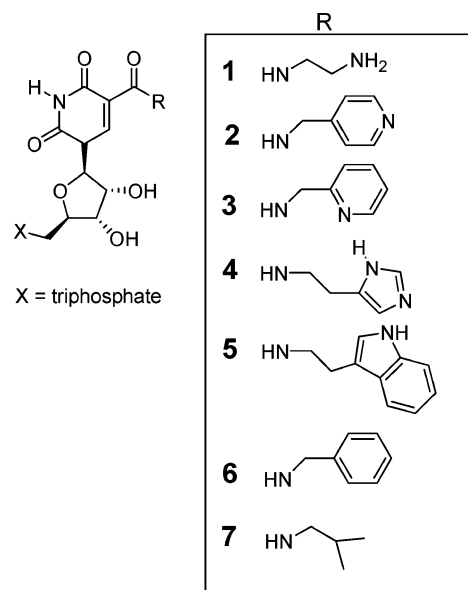

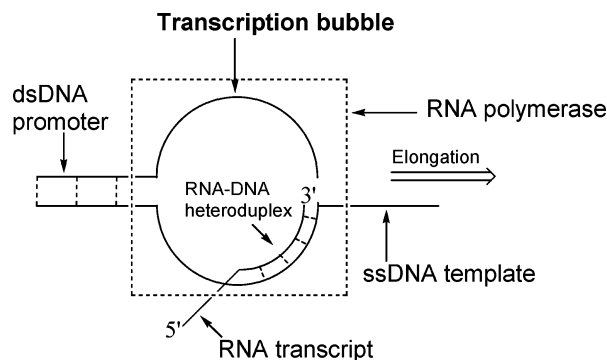**Figure 1.** UTP derivatives studied as substrates for T7 RNA polymerase.

Figure 2. Transcription bubble complex.

fications because a stable amide linkage is made to the 5-position of UTP, and essentially any primary amine can be used to impart the functional group diversity. The UTP derivatives **1**−**7** were made to determine if positively charged and hydrophobic groups could be attached to UTP and be viable T7 RNAP substrates. These types of functionality are known to be important structural components of protein enzymes, either as structural components or as participants in the chemistry at the active site. It was unclear if this range of functional groups would result in unfavorable RNA−enzyme interactions and not be tolerated as substrates for T7 RNAP.

T7 RNA Polymerase Transcription. Bacteriophage DNA-dependent T7 RNA polymerase (T7 RNAP) is one of the simplest enzymes known to synthesize RNA. It is one of several single-subunit RNA polymerases that are found in prokaryotic and eukaryotic organisms, as well as viruses. T7 RNAP was isolated in 1970 from bacteriophage T7-infected *Escherichia coli* cells.[36] T7 RNAP transcribes genes into new RNA messages from a double stranded DNA template without the use of additional transcription factors. This makes it ideally suited to an in vitro selection process. T7 RNAP was first cloned and overexpressed in *E. coli* in 1984.[37] Its availability and relative simplicity compared to those of larger multi-subunit RNA polymerases also make it attractive for the study of transcription. An integrated model of the thermodynamic, kinetic, and structural properties of the T7 RNAP transcription complex has been proposed by von Hippel.[38] The three main stages of transcription are initiation, elongation, and termination. Three crystal structures have been solved for T7 RNAP complexed to a T7 lysozyme,[39] an open 17-mer nontemplate/17-mer template promoter DNA,[40] or a 17-mer nontemplate/22-mer template with a five nucleotide overhang promoter and a short RNA transcript.[41] These crystal structures have further clarified and supported the integrated model proposed by von Hippel.

RNA polymerases initiate transcription by recognition and subsequent binding of a promoter sequence on dsDNA (Figure 2). It has been shown that the consensus promoter region of DNA must be double stranded in order for T7 RNAP to initiate transcription, while the template can be single or double stranded with no observed difference in the transcription process.[42] Once

the polymerase is bound to the promoter, the reaction is primed by a single nucleotide rather than an oligonucleotide. This single nucleotide priming is reversible and T7 RNAP transcription goes through an abortive cycling phase during the initiation step where short transcripts of 8−10 bases are formed and then dissociate from the enzyme−template complex.[43] During this abortive cycling the enzyme remains tightly bound to the promoter and the template becomes compacted into the active site. Recent crystal structures have shown that the maximum number of template bases that can be accommodated in the T7 RNAP active site is 6−9 bases, which correlates well with the length of short transcripts during the abortive cycling phase.[41] After the incorporation of 8−10 nucleotides the enzyme-template-transcript complex moves in the 3′-5′ direction along the template away from the promoter, forming a highly stable elongation complex or transcription bubble (Figure 2). Once formed, the elongation complex is highly processive and will transcribe the entire DNA template without dissociating.

Over the last few decades several groups have studied T7 RNAP transcription by kinetic analysis and mathematical modeling. From these studies the K_m values for the four naturally occurring ribonucleotides, UTP, CTP, ATP, and GTP, are reported to be in the ranges 33−60 μM, 23−81 μM, 31−76 μM, and 76−190 μM, respectively.[44−46] There are several mathematical models that have been developed to describe transcription kinetics.[44,45,47,48] Generally, these models have focused on accurately describing each stage of the transcription process separately. For example, studies focusing on initiation use short DNA templates and assume the reaction is a steady-state process with initiation being the rate-limiting step.[42,43] Other studies have used multiple long DNA templates to investigate elongation and have expressed the reaction by a simple "ping−pong" kinetic model where RNA chain initiation and termination events are positively excluded.[49] The transcription kinetics of all three stages has been modeled using a Michealis−Menten type equation.[47] A mathematical model for T7 RNAP kinetics has also been developed that takes into account essential reactant conditions, recognition sequences regulating transcription, transcript length, and sequence characteristics.[45] However, no previous study on T7 RNAP transcription has investigated the effect of modified UTP derivatives on the transcription process.

Results and Discussion

Modified UTP T7 RNAP UTP Substrate Screening. It was first of interest to determine if 5-position modified UTP derivatives with a 5-position amide linkage could be incorporated into RNA by T7 RNAP at all. Using a dsDNA template of random sequence similar to the ones used in successful in vitro selections, UTP and the UTP derivatives **1**, **2**, **4**, and 5-BU (5-benzoyl-UTP)[2] were tested under standard transcription conditions (1 mM ATP, CTP, GTP, and UTP). Figure 3 shows

(35) Ludwig, J.; Eckstein, F. *J. Org. Chem.* **1989**, *54*, 631−635.
(36) Chamberlin, M.; Mcgrath, J.; Waskell, L. *Nature* **1970**, *228*, 227-&.
(37) Davanloo, P.; Rosenberg, A. H.; Dunn, J. J.; Studier, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 2035−2039.
(38) von Hippel, P. H. *Science* **1998**, *281*, 660−665.
(39) Jeruzalmi, D.; Steitz, T. A. *EMBO J.* **1998**, *17*, 4101−4113.
(40) Cheetham, G. M. T.; Jeruzalmi, D.; Steitz, T. A. *Nature* **1999**, *399*, 80−83.
(41) Cheetham, G. M. T.; Steitz, T. A. *Science* **1999**, *286*, 2305−2309.

(42) Martin, C. T.; Coleman, J. E. *Biochemistry* **1987**, *26*, 2690−2696.
(43) Maslak, M.; Martin, C. T. *Biochemistry* **1994**, *33*, 6918−6924.
(44) Chamberlin, M.; Ring, J. *J. Biol. Chem.* **1973**, *248*, 2235−2244.
(45) Arnold, S.; Siemann, M.; Scharnweber, K.; Werner, M.; Baumann, S.; Reuss, M. *Biotechnol. Bioeng.* **2001**, *72*, 548−561.
(46) Ikeda, R. A.; Richardson, C. C. *J. Biol. Chem.* **1987**, *262*, 3790−3799.
(47) Pozhitkov, A. E.; Lavrik, I. N.; Sergeev, M. M.; Kochetkov, S. N. *Mol. Biol.* **1998**, *32*, 78−82.
(48) Young, J. S.; Ramirez, W. F.; Davis, R. H. *Biotechnol. Bioeng.* **1997**, *56*, 210−220.
(49) Rhodes, G.; Chamberlin, M. *J. Biol. Chem.* **1974**, *249*, 6675−6683.
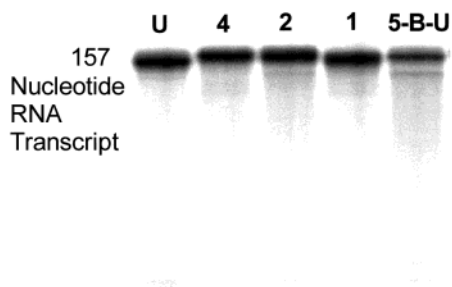
Figure 3. Phosphorimage of T7 RNAP ^{32}P body labeled transcripts from a 157 nucleotide dsDNA template with a 100 nucleotide random region. Lane **U** is the UTP control lane. Lane **4**, **2**, and **1** are transcripts using the UTP derivatives so labeled (Figure 1) in place of UTP. **5-BU** is transcript from 5-benzoyl-UTP.[2]

5'-GGG AGA CAA GAA UAA ACC UC AAU UAU GGU GUU GCC UCU UCC UAG UCG UAC UGA UGC AUC CUC UUC GAC AGG AGG CUC ACA ACA-3'

Figure 4. Phosphorimage of T7 RNAP ^{32}P body labeled transcripts from a 81 nucleotide dsDNA fixed template. Lane **U** is the UTP control lane. Lane **4**, **2**, and **1** are transcripts using the UTP derivatives so labeled (Figure 1) in place of UTP. **5-BU** is transcript from 5-benzoyl-UTP.[2]

the phosphorimage of ^{32}P body labeled RNA transcripts runs through 6% denaturing PAGE. T7 RNAP transcription using 5-BU in place of UTP had been reported previously. It should be noted that the linkage of the phenyl group in 5-BU is through a ketone linker, whereas **1**–**7** all contained an amide linkage at the 5-positon of U. In this initial screening experiment it was of interest to determine which linkage gave a greater yield of full-length transcript. For each modified uridine triphosphate, a minimum of three independent transcription reactions were run giving very reproducible levels of transcription. Compared to the UTP lane all the amide modified UTP derivatives **1**, **2**, and **4** gave good (ca. 72–105%) yields of full-length transcript, higher than the yields (41–55%) observed using 5-BU. This transcription yield is comparable to that reported previously for T7 RNAP with 5-BU.[2] All of the transcripts from the modified UTP derivatives showed a slight gel-shift retardation relative to transcript from UTP. UTP derivatives **1**, **2**, and **4** could each accommodate some degree of positive charge in the transcript, which could account for the slower gel mobility. However, 5-BU also shows a gel shift indicating that an increase in the hydrophobicity of the RNA can also give a gel shift.

From these initial transcription results we were encouraged that the amide modified UTP derivatives **1**, **2**, and **4** appeared to be suitable substrates for T7 RNAP that did not cause a significant loss in transcription yield or increase the frequency of abortive transcripts. However, T7 RNAP is an enzyme capable of giving multiple copies of transcript under the conditions used for these transcriptions, and the amount (specific activity) of the transcripts from a sequence containing a random region only depends on the average number of α-[32P]ATP incorporated. Hence the yields of full-length RNA containing the UTP derivatives relative to those obtained with UTP do not rule out a strong bias against incorporation of the derivatives, since amplification could occur with a sequence bias depleted in U.

To probe further if UTP derivatives were getting incorporated in RNA by T7 RNAP an 83 nucleotide dsDNA template of fixed sequence was prepared and used to screen UTP derivatives (Figure 4). This template codes for every triplet nucleotide sequence that contains a single uridine all linked together in one continuous sequence. It was deliberately designed to make transcription by T7 RNAP difficult. Note that even the control lane U where UTP was used in transcription showed a significant number of abortive transcripts as indicated by the faster mobility bands relative to the full-length transcript. For lane 2 using the 4-pyridyl-UTP **2**, the abortive transcripts are
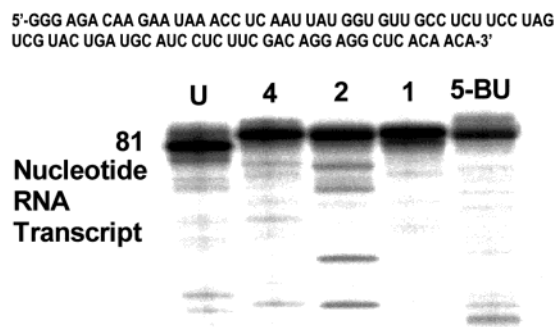
even more pronounced; however, the yield of full-length transcript is only slightly lower than that observed for the U lane. Surprisingly, lanes 4 and 1 showed no significant increase in abortive transcripts compared to that of U. In comparing the transcription results of Figures 3 and 4, it is apparent that all of the UTP derivatives are well tolerated as T7 RNAP substrates and that independent of the template the amide modifications give higher transcription yields than 5-BU. These results provided the motivation to expand the type of amide modified UTP derivatives to be studied in T7 RNAP transcription. However, it was unclear from these experiments if better yields could be obtained by increasing or lowering the concentration of the UTP derivatives or if the kinetics of transcription were different for UTP versus the UTP derivatives. To answer these questions and investigate the *relative* kinetics of **1**–**7** in T7 RNAP transcription a model system was developed.

Relative Kinetics Transcription Assay. Quantitative gel fidelity assays using polyacrylamide gel electrophoresis (PAGE) and autoradiography or phosphorimaging have been described for site-specific kinetics of DNA polymerases where relative K_m and V_{max} values can be determined by measuring integrated gel-band intensities of reaction products 5'-end labeled with [32]P.[50,51] These assays have been used to investigate the relative kinetics of several DNA polymerases where the observed integrated gel-band intensities of elongated transcript were used to approximate the relative velocity of nucleotide insertion at a specified sequence site. By measuring the relative velocities over a series of concentrations for each nucleotide, K_m and V_{max} values were obtained by fitting the relative band intensities to the Micheallis–Menten equation. A modified version of this DNA gel-shift fidelity assay was adapted for studying the relative kinetics of transcription by T7 RNAP using the seven modified UTP derivatives **1**–**7** (Figure 5).

To effectively determine how well T7 RNAP incorporates a modified nucleotide at a target site on the template, the template must be long enough so that T7 RNAP will move 3'-5' away from the promoter and form the highly stable transcription bubble. In addition the target site must be at least 10 sites downstream of the promoter so that nucleotide incorporation can be measured while the enzyme is in the elongation phase

(50) Creighton, S.; Huang, M. M.; Cai, H.; Arnheim, N.; Goodman, M. F. *J. Biol. Chem.* **1992**, *267*, 2633–2639.
(51) Boosalis, M. S.; Petruska, J.; Goodman, M. F. *J. Biol. Chem.* **1987**, *262*, 14689–14696.
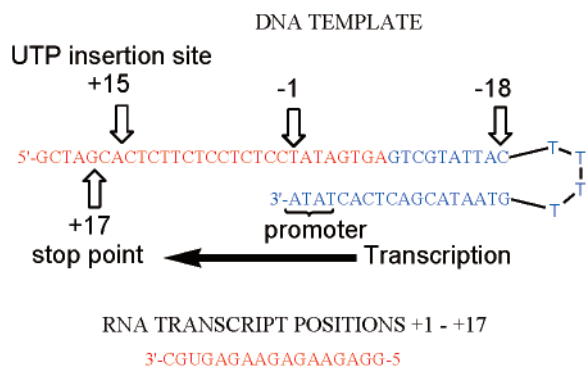
Figure 5. HPIN and UTEMP sequences anneal to form dsDNA promotor and template for T7RNAP transcritpion. Note resulting full-length transcript shown in 3′-5′ direction.
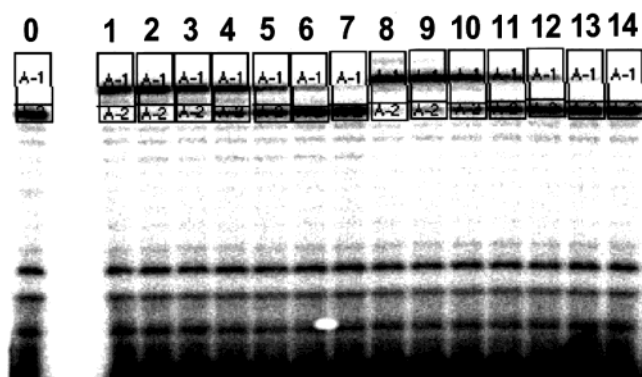

Figure 6. Phosphorimage of 5′-end labeled T7 RNAP transcripts from the template of Figure 4. Lane 0 is the transcript without UTP. Lanes 1−7 are for UTP (100−0.1 μM). Lanes 8−14 are for **6** (100−0.1 μM).

of transcription. Two synthetic DNA oligonucleotides, HPIN and UTEMP, were designed so that upon annealing a tetra-loop T hairpin was formed that contained the double stranded consensus promoter region from positions −18 to −1 with a single-stranded template extending out 21 positions from the promoter (Figure 5). This modular design is convenient in that a single promotor sequence HPIN can be combined with various templates (e.g., UTEMP) of interest. The target first site for incorporation of UTP was put at position +15 in UTEMP so that T7 RNAP would be able to form a stable elongation complex before incorporating UTP or a modified derivative into the growing transcript. The transcriptions were carried out without any CTP, so that a stop point could be inserted into the template two positions downstream from the UTP insertion site requiring T7 RNAP to incorporate a UTP, as well as the next incoming GTP before pausing. In addition, the stop point provided the opportunity to observe if any significant misin-corporation of UTP or any other NTP occurred. The transcripts were labeled by [γ-^{32}P]-GTP ensuring that only the 5′-end was labeled. The transcription products were then separated by 18% denaturing PAGE and quantitated by phosphorimaging. A representative example of the phosphorimaged PAGE is shown in Figure 6.

All determinations of relative kinetic values were performed in triplicate for **1−7** and each time with a series of UTP transcriptions as a control standard. In the example of Figure 6, UTP (lanes 1−7) and derivative **6** (lanes 8−14) are shown along with the negative control transcript minus UTP (lane 0). The A-1 boxes are the area used in phosphorimaging to determine the fraction of transcript extended beyond the stop

point +14. The A-2 boxes are the areas used in phosphor-imaging to determine the fraction of transcript stopped at +14. Without UTP (lane 0) the transcript aborts at position +14 of the template shown in Figure 5. With UTP (lanes 1−7) or **6** (lanes 8−14) the transcription extends past the UTP insertion point of position +15 including +16 and up to the stop point of +17. Note that after +14 the modified transcript (lanes 8−12) shows a significant gel shift compared to the unmodified transcript (lanes 1−6). This result is consistent with the gel shifts observed in Figures 3 and 4 for modified transcripts. Moreover, the band intensity for position +16 is not significantly more intense for UTP compared to **6** at any of the concentrations studied indicating that transcription elongation is not impeded by incorporation of the modified base. In fact at the higher concentrations of **6** it appears that some level of misincor-poration of **6** was occurring because a faint band at position +17 is observed in lane 8. Tautomerization of **6** to **8** (eq 1) may be occurring so that **6** can substitute for C at position +17.

(1)

X = triphosphate; R =HN⟨pyridyl⟩
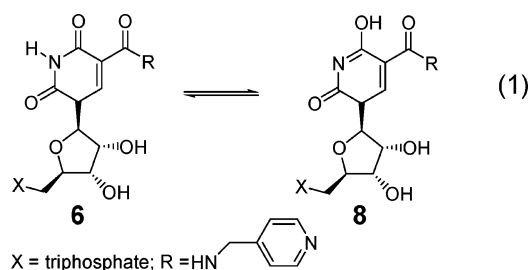
It is unclear if the misincorporation is **6** or if the incorporation of **6** at position +15 alters the transcription complex such that another nucleotide can be misincorporated at position +17. Further investigation is underway to distinguish these possibili-ties. Nevertheless, misincorporation appears to be at a low level for **6**, and similar low levels were observed for the other UTP derivatives. No such misincorporation was observed for UTP.

Relative T7 RNAP velocities were obtained by taking the ratio of phosphorimage integrated band intensities of the band corresponding to the paused transcript at position +14 (box A-2, Figure 6) over the sum of integrated band intensities of any transcript that extended out past +14 (box A-1 + box A-2, Figure 6). Since it has been shown that once T7 RNAP is bound to the promoter, it will not dissociate from the primer/template during abortive synthesis in the initiation phase or at a pausing point during elongation, keeping the polymerase and primer/template concentrations equal and high (250 nM) significantly decreasing RNAP release and turnover. Concentrations of ATP and GTP were held constant at 5 μM below their reported K_m for T7 RNAP, while changing the concentration of UTP, so that kinetic measurements reflect interactions between differ-ences in UTP derivatives and the transcription complex. These transcription conditions promote T7 RNAP into the elongation phase, so that the integrated phosphorimage PAGE gel-band intensities reflect the probability that the enzyme-primer/template complex has paused at that particular location on the template. During elongation of a growing RNA transcript, the enzyme-primer/template complex can either pause, and potentially dissociate from the growing transcript, or can add another nucleotide at each position along the template. A dark gel band represents a greater probability that the complex will pause at that particular location, while a lighter band represents a greater

	rel K_m (μM)	rel V_{max}
UTP	16	17
1	78	22
2	32	14
3	22	17
4	160	19
5	40	23
6	10	12
7	28	17

probability of polymerization. If the transcription complex is assumed to be governed only by the relative probabilities of pausing and polymerization, then a transcription complex at location +14 along the template (Figure 5) will either pause according to the probability p_{pause} or add a nucleotide to reach the target site +16 according to the probability p_{pol}.[51] These two probabilities can be expressed by the equations

$$p_{pause} = k_{pause}/(\nu_{pol} + k_{pause}) \tag{2}$$

$$p_{pol} = \nu_{pol}/(\nu_{pol} + k_{pause}) \tag{3}$$

where k_{pause} is the pausing rate of the transcription complex at the template site +14 and ν_{pol} is the rate of nucleotide addition at a target site +15 and is a function of nucleotide concentration. The integrated band intensity at position +14 (I_{+14}) and the sum of integrated band intensities at positions +15 and beyond ($I_{+15...}$) are related to the above probabilities by eq 4.

$$(I_{+15...}/I_{+14}) = (p_{pol}/p_{pause}) \tag{4}$$

Substitution of eqs 2 and 3 into eq 4 gives

$$\nu_{pol} = k_{pause} \, \rho$$

$$\rho = (I_{+15...}/I_{+14}) \tag{5}$$

The rate of nucleotide insertion, ν_{pol}, is a function of nucleotide concentration, which obeys the Michealis−Menten eq 6.

$$\nu_{pol} = V_{max} \, [\text{UTP}]/(K_m + [\text{UTP}]) \tag{6}$$

Substitution of eq 6 into eq 5 gives eq 7.

$$(I_{+15...}/I_{+14}) = \rho_{max}[\text{UTP}]/(K_m + [\text{UTP}]) \tag{7}$$

Equation 7 relates integrated gel-band intensities to nucleotide substrate concentration [UTP]. In eq 7, ρ_{max} is equal to the ratio of integrated band intensities at saturating UTP concentration, $(I_{+15...}/I_{+14})_{max}$. Therefore, $\rho_{max} = (\nu_{pol})_{max}/k_{pause}$ is relative to the maximum velocity parameter, V_{max}.

Table 1 shows the average *relative* values of K_m and V_{max} obtained from this assay for UTP and the derivatives **1**−**7**. Compared to UTP all UTP derivatives except **1** and **4** gave relative K_m and V_{max} values similar to UTP. These results are somewhat surprising, since **1** is expected to be positively charged and hydrophilic, whereas **2**, **3**, **5**, **6**, and **7** all introduce varying degrees of increased hydrophobicity in the UTP substrate. UTP derivative **4** may also be positively charged and has some hydrophobic character. It is tempting to speculate that the relative K_m of **4** is higher because of steric interactions at the

T7 RNAP active site, since groups of similar size in **2**, **3**, and **6** but attached by one fewer CH_2 linker gave significantly lower K_m values. However, **5** does not show this increase in relative K_m and its modification is attached by the same linker as **4** and is even larger. It is possible that the imidazole group in **4** and the amino group in **1** both require deprotonation prior to entering the T7 RNAP active site, if for example the counterion cannot be accommodated or these groups stabilize a conformation of the triphosphate anion incompatible with transcription. Nevertheless, the relative V_{max} of **4** and **1** are essentially unaffected compared to UTP or the other UTP derivatives. It appears that all of these UTP derivatives are suitable substrates for T7 RNAP and that with the exception of **4**, there is no need to use elevated concentrations of these derivatives relative to UTP.

Conclusions

UTP derivatives modified at the 5-position with amide linkers are tolerated as substrates by T7 RNAP. Compared to the previously reported 5-BU UTP derivative which has a phenyl group connected by a ketone linker, transcription yields were higher for all the amide linked derivatives. Transcription yields of T7 RNAP using random RNA sequences gave good yields of a full-length transcript and few aborts when using the UTP derivatives **1**, **2**, and **4** compared to UTP. A challenging sequence that caused T7 RNAP abortion of transcription using UTP also gave good yields of full length transcript using **1** and **4** with a slight increase relative to UTP in the number of abortive transcripts for **2**. Measuring the relative kinetics of an expanded group of UTP derivatives gave more insight into the tolerance of T7 RNAP for these modified substrates.

Using a designed template under transcription conditions that favor association of T7 RNAP to the transcription complex, while varying the concentration of UTP, the relative K_m and V_{max} values were determined for the amide linked 5-position-modifed UTP derivatives **1**−**7**. Surprisingly, the size of the modification had less of an affect on the relative K_m than the potential to bear a positive charge. Both **1** and **4** showed a significant increase in K_m compared to UTP or the other derivatives. Perhaps more surprising, the relative V_{max} values were all similar to UTP independent of the structure of the modification. From the phosphorimage band intensities of the transcripts from the kinetic assay it was clear that no significant pause in elongation occurred with any of the UTP derivatives compared to UTP. It appears that incorporation of these amide linked UTP derivatives does not impede elongation and that the modified transcript does not significantly alter the transcription bubble in a way that slows polymerization.

It is now known that **1**−**7** can be used as substrates for T7 RNAP and could prove useful for future RNA in vitro selections. Even charged, hydrophobic, alkyl, and large aryl groups at the 5-position of UTP were well tolerated. Furthermore, under the typical conditions used for T7 RNAP transcription (1 mM ATP, CTP, GTP, UTP) **1**−**7** would all be above saturation of the polymerase. However, trace levels of misincorporation of **1**−**7** in place of C were observed at concentrations of 1 mM using the UTEMP DNA template. To avoid this potential for misincorporation, transcription should be performed with these UTP derivatives at 100 μM. Further research is underway to determine the extent of misincorporation when using these amide linked UTP derivatives.

Experimental Section

General. UTP derivatives **1**−**7** were prepared by carboxyamidation[24] followed by triphosphate synthesis as described previously.[35] The 5′-triphosphates **1**−**7** were purified by DEAE Sephadex anion exchange column (0.05−2.0 M triethylammonium bicarbonate pH 8.0) followed by reverse-phase HPLC (C18 column, 0−100% CH₃CN in 0.10 M triethylammonium. The nucleoside precursors and triphosphates were checked for purity by analytical HPLC (C18 column, gradient as above), ^{31}P and ^{1}H NMR, and mass spectrometry.[52] The triphosphate solutions were quantitated on the basis of their UV absorbance at λ_{max}.[52]

Oligonucleotides were synthesized on an ABI model 392 DNA synthesizer or obtained unpurified from Operon Technologies, Inc.. The 100N random region was synthesized with an equimolar mix of cyanoethyl phosphoramidites prepared by Glen Research. Crude ssDNA was purified by denaturing polyacrylamide gel electrophoresis (6% or 8% acrylamide, 19:1 acrylamide/bisacrylamide) and visualized by UV shadowing. The purified oligonucleotides were eluted from the crushed gel slices with 2 mM EDTA, filtered through 0.45 μm cellulose acetate, desalted on sephadex G-25, and then lyophilized to dryness and resuspended in deionized H₂O. The transcription template stock solution was prepared by annealing 10 μM each of HPIN and UTEMP (Figure 5) in 100 mM NaCl, 10 mM Tris·HCl pH 7.5 at 95 °C for 5 min, followed by slow-cooling to 25 °C over 20 min. The dsDNA templates used for the T7 RNAP transcription experiments (Figures 3 and 4) were obtained by PCR amplification of the corresponding ssDNA with fixed-sequence primers containing the T7 RNAP promotor necessary for transcription. Nucleotide triphosphates (100 mM) were obtained from Pharmacia-Biotech. T7 RNA polymerase (14.7 μM, ca. 90 U/μL) was obtained from Enzyco (Denver, CO). The manufacturers' supplied buffers were used for all enzymes unless otherwise noted. γ-[^{32}P]GTP (6000 Ci/mmol), γ-[^{32}P]ATP (6000 Ci/mmol), and α-[^{32}P]ATP (800 Ci/mmol) were from Amersham. PAGE images were obtained with a Cyclone phosphorimager. Integrated band intensities of the transcription products were obtained using OptiQuant software.

Transcriptions Screening UTP Derivatives. Transcriptions on the 100N random template (Figure 3) and the fixed sequence template (Figure 4) were carried out in 100 μL of a buffer consisting of 40 mM Tris·HCl pH 8, 12 mM MgCl₂, 1 mM spermidine, 5 mM dithiothreitol (DTT), 4% glycerol, and 0.01% Triton-X in the presence of 1 mM

each ATP, CTP, GTP, and UTP at 250 nM dsDNA template, 20 mM GMP, 10 μCi α-[^{32}P]ATP, 50 units of RNAse inhibitor, with 0.88 μM T7 RNA polymerase. The transcriptions were run at 37 °C for 20 h and stopped by the addition of 10 μL of 100 mM EDTA. After desalting the transcriptions on 10K molecular weight cutoff filters, the RNA was loaded on either 6% or 8% denaturing acrylamide gels, preheated to 45−50 °C. The gels were dried in vacuo and phosphorimaged to quantitate the full-length RNA.

Relative T7 RNAP Transcription Kinetics Assay. The relative kinetic assay transcription reactions were carried out in 15 μL of 40 mM Tris(pH 7.9), 6 mM MCl2, 2 mM spermidine, 10 mM NaCl, 5 μM ATP(Sigma-Aldrich), 5 μM GTP(Sigma-Aldrich), 250 nM template, and 250 nM T7 RNAP (Promega). The concentration of UTP and the UTP derivatives **1**−**7** was varied over a range of 0.1 μM to 1mM. [γ-^{32}P]GTP(ICN)was added to approximately 0.5 μCi/μL. A master mix was prepared by adding HPIN and UTEMP (Figure 5) to a sterile Epindorf tube and incubating at 25 °C for 5 min before adding ATP, GTP, [γ-32P]GTP, 5x transcription buffer (Tris, MgCl2, spermidine, NaCl), and RNAse-free water to bring the concentrations of each reagent to those listed above. Aliquots of the master mix were added to separate Epindorf tubes for each reaction. UTP or a UTP derivative was then added separately to each tube. For each set of reactions performed with a UTP derivative, an identical set of control reactions was run simultaneously with UTP. The reaction was initiated by the addition of T7 RNAP and incubated at 37 °C for 10 min. Each transcription was quenched by adding 15 μL of formamide containing 20 mM EDTA, 0.05% xylene cyanol, and 0.05% bromophenol blue. The samples were heated for 2 min at 70 °C and analyzed on a preheated (45−50 °C) 20% denaturing PAGE (20 × 20 × 0.04 cm³). After the gels were dried in vacuo, phosphorimaging was used to quantitate the relative band intensities.

Supporting Information Available: Additional information on kinetic data and plots (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA049009H

---

(52) Supporting Information.