Preparation of probe-modified RNA with 5-mercapto-UTP for analysis of protein–RNA interactions

Bokang He[1], Daniel L. Riggs[1] and Michelle M. Hanna[1,2,*

Departments of [1]Botany/Microbiology and [2]Chemistry/Biochemistry, The University of Oklahoma, Norman, OK 73019, USA

ABSTRACT

We report a modified synthesis for 5-mercapto-UTP (5-SH-UTP) and its use for analysis of protein–RNA interactions utilizing *Escherichia coli* and T7 RNA polymerases and yeast RNA polymerases I and III. 5-SH-UTP did not affect transcriptional pausing, Rho-independent termination or recognition of the *E.coli* transcription complex by NusA. RNA containing 5-SH-UMP did not crosslink to polymerase when irradiation was at 302 or 337 nm. Transcription complexes containing RNA substituted with 5-SH-UMP were post-transcriptionally modified to attach a photocrosslinking group to thiol-tagged nucleotides in the RNA on the surface of the polymerase or free in solution. The pK_a for 5-SH-UTP was determined to be 5.6, so modification of the thiol groups in the RNA with *p*-azidophenacyl bromide could be carried out at pH 7. Addition of the transcription termination factor Rho, a RNA binding protein, to *E.coli* transcription complexes resulted in RNA crosslinking to Rho and to the β and β' subunits of polymerase. Using 5-SH-UTP, one can distinguish RNA binding domains on the surface of RNA polymerases or other RNA binding proteins from those buried within the protein.

INTRODUCTION

Almost all RNA synthesis is catalyzed by the enzyme RNA polymerase, which forms an RNA molecule, complementary to one strand of a DNA template, by polymerization of ribonucleoside triphosphates (1–4). The complexity of RNA polymerases from different organisms varies greatly, ranging from single polypeptides, in bacteriophages such as T7, to more than 10 polypeptides, most of which have not yet been isolated, in eucaryotic polymerases (3,5). Much work is now focused on characterization of the structure of various polymerases, including identification of the specific protein–nucleic acid interactions involved in promoter recognition and RNA synthesis (6,7). By far the best characterized RNA polymerase is that from the bacterium *Escherichia coli*. Detailed molecular analysis of the transcription complex and how this complex is modified by various viral proteins (8–10) can now be achieved, because the polymerase subunits and their genes have been isolated.

For several years our work has focused on analysis of the protein–RNA contacts in *E.coli* initiation, elongation and anti-termination complexes. We have used a variety of photocrosslinking nucleotide analogs, which can be incorporated into the nascent RNA *in vitro*, to identify the different subunits of *E.coli* RNA polymerase which are contacted by the RNA during transcription (11–16). By using several different photocrosslinking nucleotide analogs, it has been possible to identify the polymerase subunits which are contacted by the 5'-end, the 3'-end and internal regions of the nascent RNA in active transcription complexes. While these studies have allowed identification of the polymerase subunits which are contacted by the nascent RNA during transcription, they have not provided information on the nature of the RNA binding regions on these subunits. RNA crosslinking of a subunit merely identifies that subunit as being on the path of the nascent RNA during transcription, but gives no information as to whether this path is on the surface of the enzyme or if a specific, solvent-inaccessible, RNA binding tunnel exists. While the existence of such an RNA binding tunnel is now part of several models for the *E.coli* transcription complex (for reviews see 17–19), there currently exist no definitive studies which clearly establish the existence or address the nature of such a tunnel.

We report here a method to determine which subunits of RNA polymerase interact with the RNA on the surface of the enzyme and to distinguish between interactions with these surface binding sites and interactions with subunits that occur when the RNA is buried within the transcription complex. This can be done by utilizing the two analogs, 5-mercapto-UTP (5-SH-UTP; Fig. 1A, II) and 5-[4-(azidophenacyl)thio]-UTP (5-APAS-UTP; Fig. 1A, III). We have shown previously that 5-APAS-UTP is a substrate for *E.coli* RNA polymerase (12). The RNA formed appears to be threaded properly through the RNA binding domains of the polymerase, because transcriptional pausing and termination are unaffected and the transcription factor NusA still interacts with the RNA in the transcription complex (13). We have now characterized the substrate properties of 5-SH-UTP, which is the immediate precursor in the synthesis of 5-APAS-UTP. We show that it is a substrate for *E.coli* and bacteriophage T7 RNA
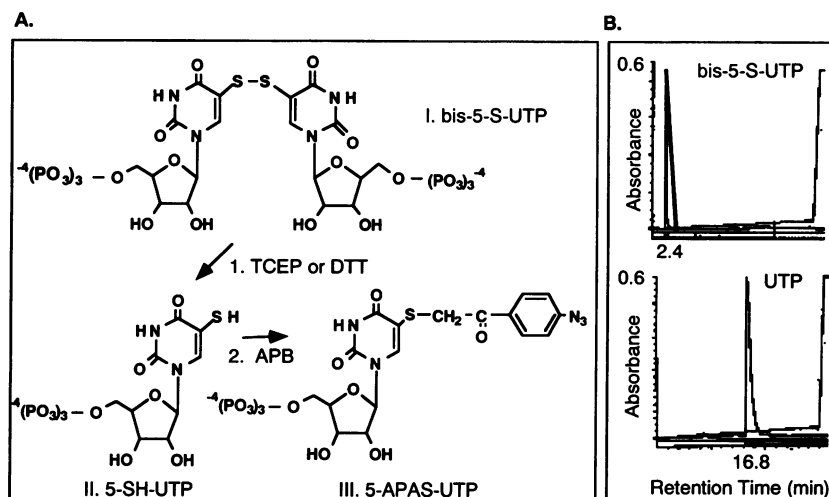
Figure 1. Synthesis of 5-SH-UTP and 5-APAS-UTP. (A) 5-SH-UTP (II) was synthesized from UTP via a dinucleoside hexaphosphate intermediate (bis-5-S-UTP, I), in which two 5-SH-UTP molecules were joined at the 5 position of the base with a disulfide bond. Upon reduction of this disulfide bond with TCEP or DTT, 5-SH-UTP was formed. 5-APAS-UTP (III) was formed by attachment of APB, to the mercapto group of 5-SH-UTP (II). (B) The HPLC chromatograms of the purified bis-5-S-UTP and UTP are shown. Bis-5-S-UTP elutes at 2.4 min and it is free of contaminating UTP, which elutes at 16.8 min.

polymerases and yeast RNA polymerases I and III. We have incorporated this analog into RNA during transcription and then modified the RNA with azidophenacyl bromide (APB) post-transcriptionally, while it is still part of the ternary transcription complex. Only nucleotides which are on the surface of the enzyme or free in solution will be accessible to APB.

Although we have previously published the synthesis of 5-SH-UTP (20,21), it is extremely difficult to synthesize and in the past has often given poor yields. We have now improved upon the procedure and here we present modifications to the method and describe the changes which have allowed us to significantly increase the yield and purity.

MATERIALS AND METHODS

Materials and buffers

Ultrapure ribonucleoside triphosphates (NTPs) were purchased from Pharmacia LKB Biotechnology Inc. ApApC, ApU and GpG and APB were purchased from Sigma. Nitrocellulose transfer membrane (0.45 μm) was the product of Micron Separations Inc. Buffer A: 30 mM Tris–OAc, pH 8.0, 40 mM KOAc, 0.5 mM Na_2EDTA, 1 mM dithiotheitol (DTT), 50 μg/ml acetylated bovine serum albumin (BSA); buffer B: 40 mM Tris–HCl, pH 7.6, 15 mM $MgCl_2$, 10 mM β-mercaptoethanol, 50 μg/ml acetylated BSA; buffer C: 30 mM Tris–HCl, pH 7.0, 10 mM KCl, 0.5 mM Na_2EDTA, 50 μg/ml acetylated BSA; buffer D: 7 M urea, 10 mM Na_2EDTA, 0.02% (w/v) bromophenol blue, 0.02% (w/v) xylene cyanol; buffer E: 100 mM Tris–HCl, pH 8.0, 200 mM DTT, 4% (w/v) SDS, 0.02% (w/v) bromophenol blue, 20% (v/v) glycerol; buffer F: 0.3% (w/v) Tris base, 1.44% (w/v) glycine, 15% (v/v) methanol, 0.01% (w/v) SDS, pH 8.3; buffer G: 2% (w/v) Na_3-citrate, 2% (w/v) $FeSO_3 \cdot 7H_2O$, 0.1% (w/v) $AgNO_3$.

Enzymes and plasmids

RNA polymerase was purified from *E.coli* MRE 600 cells (Grain Processing Co., Muscatine, IL) using published procedures

(22,23). Restriction enzymes were the products of Promega. T1 RNase was purchased from Sigma. Plasmid pKK34-121 contains the *E.coli* *rrn*B P1 and P2 promoters (24). T7 RNA polymerase and plasmid pLM45, containing the phage T7 φ10 promoter, were kindly provided by Dr William McAllister (SUNY). Plasmid pAR1707 contains the A1 promoter from bacteriophage T7, which is a strong *E.coli* RNA polymerase promoter, and the Rho-independent early terminator t_e. Plasmid pBHE-179 contains the bacteriophage λ $P_{R'}$ promoter (an *E.coli* RNA polymerase promoter), with an *Eco*RI site at position +179 relative to the start of transcription, cloned into the multiple cloning site of pBluescript KSII (–) (Stratagene). Gln111, which binds strongly to an *Eco*RI site but does not cleave the DNA, was purified as described (25,26). Transcription factor NusA was made as a fusion protein in our laboratory (27). Termination factor Rho was kindly provided by Dr D. Bear (University of New Mexico).

Synthesis of 5-SH-UTP

5,5″-Dithiobis(uridine 5′-triphosphate) (bis-5-S-UTP; Fig. 1A, I) was synthesized using a modification of a published method (20,21). Fresh reagents must be used. UTP (1.0 g, 2 mmol) was suspended in 75 ml dimethylacetamide at 4°C. Methylhypobromite was prepared in 38 ml anhydrous methanol from 14 g (50 mmol) silver carbonate and 2.8 ml (50 mmol) bromine. Sodium disulfide was prepared from 13 g (48 mmol) sodium sulfide nonahydrate and 1.6 g (48 mmol) sulfur in 75 ml 100% boiling ethanol. The methylhypobromite was added to the suspension of UTP-dimethylacetamide, followed by addition of 13 g (40 mmol) freshly prepared sodium disulfide. The reaction mixture was stirred at room temperature for 24 h. Separation by anion exchange column chromatography (28) gave a 42% yield of bis-5-S-UTP. The purity of the bis-5-S-UTP was determined by HPLC analysis on a Beckman Ultrasphere ODS column (4.6 × 25 cm). The samples were eluted at 1 ml/min with a 30 min linear gradient of 0–5% (v/v) acetonitrile in 50 mM triethylammonium bicarbonate (pH 8). The retention time for bis-5-S-UTP is 2.6 min and the retention time for UTP is 16.7 min. The bis-5-S-UTP is
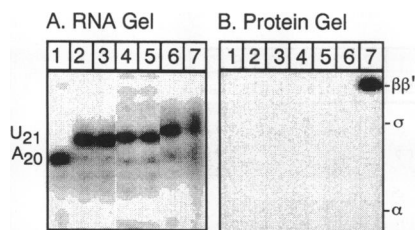
Figure 2. Specificity of incorporation of 5-SH-UTP with *E.coli* RNA polymerase and evaluation of its photocrosslinking properties. Paused transcription complexes which contained RNA 20 nt long (designated A$_{20}$) were prepared from the T7 A1 promoter with ApU, [α-^{32}P]GTP, CTP and ATP. The RNA sequence encoded after position +20 is UAC. A$_{20}$ complexes were isolated and then incubated with buffer (lane 1), 1 μM UTP (lanes 2 and 3), 1 μM 5-SH-UTP (lanes 4 and 5) or 10 μM 5-APAS-UTP (lanes 6 and 7) to allow elongation to position 21. Half of the reactions were irradiated with a 302 nm light source (lanes 3, 5 and 7) to crosslink the RNA to *E.coli* RNA polymerase subunits. (**A**) Autoradiogram of the RNA gel. The position of the U-21 RNA made with UTP is indicated. RNA containing 5-SH-UMP or 5-APAS-UMP has a somewhat slower mobility. (**B**) Autoradiogram of the SDS–PAGE gel on which samples were analyzed. Positions of the RNA polymerase subunits, as determined by silver staining of the gel prior to autoradiography, are shown.

completely free of contaminating UTP (Fig. 1B). 5-SH-UTP (Fig. 1A, II) was made by reducing bis-5-S-UTP with DTT at pH 8.0. The reaction contained 4 mM bis-5-S-UTP, 10 mM Tris–HCl (pH 8.0) and 10 mM DTT and was incubated at room temperature for 10 min. The concentration of 5-SH-UTP was determined using ε_{336} = 8000/cm/M (29). 5-SH-UTP was prepared by reduction of bis-5-S-UTP immediately before use. Even overnight storage at –20°C resulted in complete re-oxidation of the 5-SH-UTP to bis-5-S-UTP, but this sample could be reduced a second time by addition of fresh DTT. Bis-5-S-UTP can also be be reduced with (tris-2-carboxyethyl)phosphine hydrocloride (TCEP) (30).

5-APAS-UTP (Fig. 1A, III) was synthesized from bis-5-S-UTP and purified as previously described (20). 4-Thio-UTP (4-S-UTP) was synthesized from 4-S-UDP and purified as previously described (31). Transcription reactions with both analogs were carried out in reduced light, as were all reactions involving APB.

Specificity of incorporation of 5-SH-UTP with *E.coli* RNA polymerase and evaluation of its photocrosslinking properties

Transcription was started by mixing 10 nM plasmid pAR1707, 30 nM *E.coli* RNA polymerase, 25 μM ApU and 5 mM Mg(OAc)$_2$ in buffer A. The mixture was pre-incubated at 37°C for 5 min, then 1 μM ATP, 1 μM CTP, 330 nM [α-^{32}P]GTP (6.7 × 10^6 d.p.m./pmol) and 10 μg/ml rifampicin were added to the reaction. This resulted in formation of A$_{20}$ complexes, which are ternary transcription complexes containing RNA 20 nt long. The RNA sequence from this promoter is 5'-AUCGAGAGGGA-CACGGCGAA$_{20}$UAC. Therefore, initiation of transcription in the presence of ApU, CTP, GTP and ATP causes the RNA polymerase to pause after incorporation of the first 20 nt in the absence of UTP. To test the specificity of incorporation of 5-SH-UTP by elongation of the 20mer RNA to the 21mer RNA, A$_{20}$ complexes were separated from unincorporated nucleotides by gel exclusion chromatography, as previously described (12). The 20mer RNA was then elongated to the 21mer by addition of

UTP (1 μM), 5-SH-UTP (1 μM) or 5-APAS-UTP (20 μM). RNA was analyzed by electrophoresis on a 20% polyacrylamide–urea gel (Fig. 2A).

To compare the photocrosslinking properties of each of the UTP analogs and UTP, transcription complexes containing the 21mer RNAs were placed in clear plastic microcentrifuge tubes ~1.5 cm above a 302 nm light source (Spectroline model XX-15B, 1800 μW/cm^2 at 15 cm) and irradiated for 2 min at room temperature. DTT was added to each sample and then the proteins were analyzed on a 10% SDS–PAGE gel, followed by silver staining and autoradiography of the gel (Fig. 2B).

Transcription properties of *E.coli* RNA polymerase as a function of pH

The pK_a for the thiol group on 5-SH-UTP was estimated by analysis of the absorption spectrum in the pH range 5–8 (Fig. 3A). For transcription over this pH range, the pH of transcription buffer A was varied from 5.0 to 8.0 by replacing Tris–HCl with phosphate. Reactions contained 25 nM plasmid pKK34-121, 100 nM RNA polymerase, 200 μM ATP, 200 μM CTP, 200 μM UTP and 20 μM GTP and they were pre-incubated at 37°C for 10 min. After pre-incubation, 5 mM Mg(OAc)$_2$, 10 μg/ml rifampicin and 165 nM [α-^{32}P]GTP (6.7 × 10^6 d.p.m./pmol) were added and the reaction was incubated for an additional 10 min. The RNA was isolated and analyzed as described above (Fig. 3B).

Effect of 5-SH-UTP on *E.coli* RNA polymerase pausing and recognition of the transcription complex by NusA

Transcription was initiated with 30 nM *E.coli* RNA polymerase, 10 nM DNA template (pAR1707), 25 μM ApU and 5 mM Mg(OAc)$_2$ in buffer A with or without NusA (90 nM). After incubation at 37°C for 5 min, 1 μM ATP, 1 μM CTP, 660 nM [α-^{32}P]GTP (6.7 × 10^6 d.p.m./pmol) and 10 μg/ml rifampicin were added and the reaction was incubated at 37°C for an additional 2 min. A nucleotide chase mix containing 20 μM ATP, CTP, GTP and 5 mM Mg(OAc)$_2$ in buffer A with either 20 μM UTP or 100 μM 5-SH-UTP was then added to the reaction. Samples were withdrawn from the reaction at various times (0–15 min) for analysis of the RNA. Samples were transferred from the reaction mixture directly into the same volume of buffer E to stop transcription. The RNA was analyzed by electrophoresis and autoradiography (Fig. 4).

Transcription with T7 RNA polymerase

Transcription was initiated with 40 nM T7 RNA polymerase, 20 nM plasmid pLM45, 40 μM GpG, 100 μM ATP, 100 μM CTP, 10 μM [α-^{32}P]GTP (10^5 d.p.m./pmol) for 5 min at 37°C in buffer B. UTP or 5-SH-UTP (100 μM) and heparin (80 μg/ml) were then added, followed by incubation of the reactions at 37°C for an additional 10 min. The reactions were extracted one time with phenol:chloroform:isoamyl alcohol (25:24:1). The RNA was precipitated in ethanol, resuspended in buffer E and heated at 90°C for 2 min before electrophoresis (Fig. 5A).

Transcription with yeast RNA polymerases I and III

The protocols for preparation of the transcription extracts, the templates used and the transcription assays were those described by Riggs *et al.* (32). RNA polymerase I transcription was
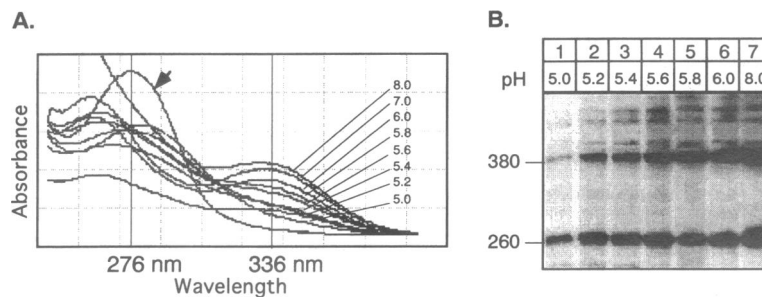
Figure 3. Properties of 5-SH-UTP and *E.coli* RNA polymerase as a function of pH. (**A**) Absorption characteristics of 5-SH-UTP as a function of pH. The absorption spectrum of 5-SH-UTP was recorded in solutions in which the pH was varied from 5.0 to 8.0. The absorption maximum characteristic of the 5-mercapto group is at 336 nm when the thiol is ionized. The absorption maximum of bis-5-S-UTP, which reforms upon oxidation of 5-SH-UTP, is at 273–276 nm. The absorption spectrum of bis-5-S-UTP is indicated with an arrow. 5-SH-UTP is 50% ionized at pH 5.6. (**B**) Transcription properties of *E.coli* RNA polymerase as a function of pH. The effect of pH on *E.coli* RNA polymerase was evaluated using transcription from the *E.coli* *rrn*BP1 and P2 promoters. The lengths of the RNA species produced by Rho-independent termination at the terminator t_1 are 260 and 380 nt.
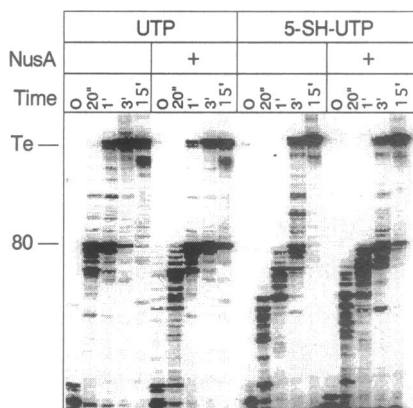

Figure 4. Comparison of *E.coli* RNA polymerase pausing with UTP and 5-SH-UTP. Transcription complexes paused at position +20 were prepared from the T7 A1 promoter. The RNA was then elongated in the presence of 20 μM ATP, GTP, CTP and either 20 μM UTP or 100 μM 5-SH-UTP, with and without NusA present. Aliquots were withdrawn at the indicated times and the RNA was analyzed by denaturing PAGE. The major paused species from this template occurs at +80 and this pause is enhanced by NusA. The terminator t_e is a Rho-independent terminator.

analyzed using the low salt pellet fraction, while RNA polymerase III transcription was analyzed using the low salt supernatant fraction. The RNA polymerase I template used was pDR10 linearized with *Eco*RV and the RNA polymerase III template was the 5S rDNA gene-containing plasmid pBB111R. The only modification of the transcription assay was that the total volume was reduced to 20 μl and the nucleotide substrates were adjusted as described in the figure legends (Fig. 5B and C).

Post-transcriptional modification of RNA in transcription complexes with APB

RNA containing 5-SH-UMP was made with *E.coli* RNA polymerase. The DNA template was a PCR fragment made from the plasmid pBHE-179 using the T3 and T7 primers from Stratagene. Transcription was started by mixing 40 nM *E.coli* RNA polymerase, 10 nM DNA template, 1 μM ApApC, 0.1 mM ATP, 0.1 mM CTP, 10 μM GTP and 20 nM Gln111 in buffer C. After a 10 min incubation at 37°C, 5 mM MgCl₂, 100 μM
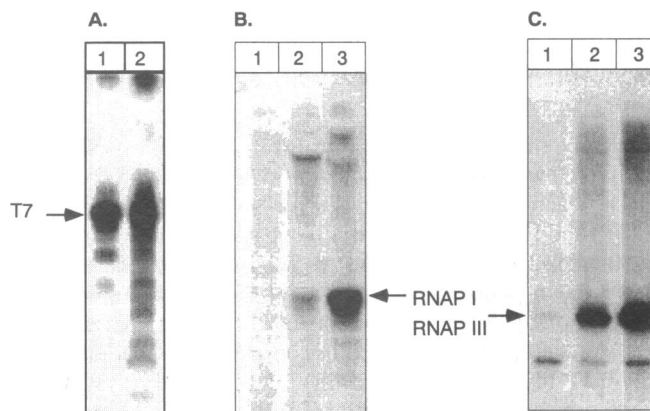

Figure 5. Substrate properties of 5-SH-UTP with other RNA polymerases. (**A**) Transcription was carried out with bacteriophage T7 RNA polymerase in the presence of ATP, CTP, GTP and 100 μM UTP (lane 1) or 100 μM 5-SH-UTP (lane 2). (**B** and **C**) Transcription assays were performed with whole cell extracts from yeast which contained all of the factors necessary for specific transcription by RNA polymerase I (B) or RNA polymerase III (C). Transcription was directed by a linearized rDNA or 5S template in the presence of ATP, CTP (200 μM each) and GTP (15 μM), in the absence of UTP or analog (lane 1) or in the presence of 5-SH-UTP (500 μM, lane 2) or UTP (200 μM, lane 3). The positions of the specific transcripts are indicated by the arrows.

5-SH-UTP or 100 μM UTP, 0.33 μM [α-^{32}P]GTP (6.7 × 10⁶ d.p.m./pmol) and 10 μg/ml rifampicin were added to the reaction, followed by a 15 min incubation at 37°C. The RNA produced was 165 nt long and was still associated with the transcription complex, because Gln111 bound at the *Eco*RI site at +179 functions as a transcription block, with the 3'-end of the resulting RNA being 14 nt upstream from the *Eco*RI site (33). There was no released RNA in this reaction (not shown).

Reactions were split and to half was added APB, which was dissolved in dimethyl sulfate, to a final concentration of 5 mM. After incubation at room temperature for 2 h, the reaction was passed through a Sephadex G-50 spun column pre-equilibrated in buffer C without BSA. Rho protein (200 nM) was then added, followed by incubation at 37°C for 10 min. Half of the reaction was then irradiated at 302 nm for 2 min (12) and half was kept in the dark at room temperature. DTT (100 mM) was then added to
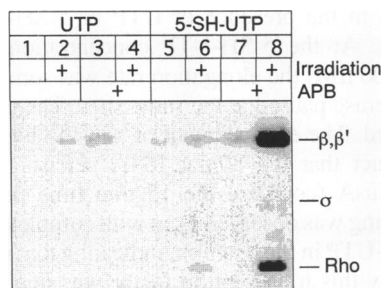
Figure 6. Specificity of crosslinking by RNA containing 5-SH-UMP post-transcriptionally modified with APB. Transcription was carried out with *E.coli* RNA polymerase in the presence of UTP or 5-SH-UTP under conditions which yielded ternary transcription complexes containing RNA 165 nt long and no released RNA. Samples were divided and half of each was treated with APB at pH 7.0 to allow modification of surface exposed thiol residues. Transcription complexes were purified by gel exclusion chromatography and the RNA binding termination factor Rho was then added to each reaction. The reactions were irradiated for 2 min at 302 nm to allow crosslinking of the RNA to RNA polymerase subunits or Rho. Proteins were then separated by SDS–PAGE. Proteins were electrophoretically transferred to a nitrocellulose membrane, which was then silver stained to locate proteins. The nitrocellulose filter was then exposed to X-ray film to allow identification of proteins radioactively labeled by crosslinking to the RNA.

Figure 7. Evaluation of different UTP analogs as photocrosslinking probes. Ternary transcription complexes were formed with *E.coli* RNA polymerase which contained RNA 165 nt long. The RNA was synthesized in the presence of ATP, CTP, GTP and either 4-thio-UTP, UTP or 5-SH-UTP. Complexes were either maintained in the dark as controls or were irradiated with a 302 nm light source or a 337 nm nitrogen laser for 1 or 10 min. Reactions were then loaded onto an SDS–PAGE gel and, after electrophoresis, the proteins were transferred to a nitrocellulose membrane. The membrane was silver stained and subjected to autoradiography. The positions of the β and/or β′ subunits of RNA polymerase are indicated.

all reactions. The control and the irradiated aliquots were then split in half for RNA analysis on an 8% polyacrylamide–urea gel in buffer D and for protein analysis. The RNA in the samples to be run on protein gels was digested with ribonuclease T1 at 37°C for 30 min and then an equal volume of buffer E was added. Proteins were analyzed on a 6% Tricine–SDS–polyacrylamide gel (34). After electrophoresis, proteins were electrophoretically transferred to a nitrocellulose membrane (0.45 µm; Micron Separations Inc.) using buffer F (150 mA, 6 h). The membrane was stained with buffer G and then exposed to X-ray film at –80°C (Fig. 6).

Evaluation of 5-SH-UTP as a photoreactive nucleotide analog

Transcription was initiated in the dark by incubation of 10 nM DNA (PCR fragment of pBHE-179), 10 µM GTP, 100 µM ATP and CTP, 40 nM *E.coli* RNA polymerase and 20 nM Gln111 in buffer C at 37°C for 10 min. The reaction was then split into three aliquots, and 0.33 µM [α-^{32}P]GTP (7 × 10^6 d.p.m./pmol), 10 µg/ml rifampicin, 5 mM MgCl$_2$ and 100 µM either 4-thio-UTP, UTP or 5-SH-UTP were added to each aliquot. Reactions were incubated at 37°C for 15 min and transcription was then stopped by addition of 10 mM EDTA. Transcription complexes were purified by gel exclusion chromatography on a Sephadex G-50 spun column pre-equilibrated in buffer C without BSA. One fifth of each reaction was removed and kept in the dark. The remaining reactions were again split in half and half of each reaction was irradiated with the 302 nm light source, as described above. The other half was irradiated at 337 nm with a nitrogen laser (model LN 1000; Photochemical Research Associates Inc.). Samples were removed after 1 and 10 min irradiation. DTT (to 100 mM) was added to each sample and they were analyzed on an RNA (not shown) or protein gel (Fig. 7), as described above. The samples for protein analysis were treated with ribonuclease T1 at 37°C for 30 min before electrophoresis. After electrophoresis, proteins
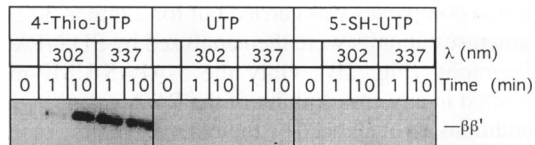
were electrophoretically transferred to a nitrocellulose membrane (0.45 µm; Micron Separations Inc.) using buffer F. The membrane was silver stained with buffer G and then exposed to X-ray film at –80°C.

RESULTS AND CONCLUSIONS

Synthesis and characterization of 5-SH-UTP

The synthesis of 5-SH-UTP has been described previously as part of the synthetic pathway for 5-APAS-UTP (Fig. 1A). This synthesis contains several steps at which slight deviations from the optimal protocol will cause the reaction to fail. The modifications that we report here have allowed us to reproducibly obtain much higher yields of 5-SH-UTP than with the earlier procedure. The 5-SH-UTP, generated by reduction of bis-5-S-UTP, contains no detectable UTP, as judged by analytical HPLC (Fig. 1B). It is critical that all reagents be fresh (new bottles) and that the synthesis be carried through to as far as the synthesis of bis-5-S-UTP before stopping. The earlier synthesis involved reaction of sodium disulfide with methylhypobromite-UTP suspension at 4°C. We have now evaluated the effect of reaction times and temperatures during this step and have found that the yield is significantly increased when the reaction is done at room temperature for at least 24 h.

Specificity of incorporation of 5-SH-UTP with *E.coli* RNA polymerase

To determine if 5-SH-UTP could specifically replace UTP during transcription with *E.coli* RNA polymerase, transcription complexes containing RNA 20 nt long prepared from the T7 A1 promoter were isolated. This is a strong *E.coli* RNA polymerase promoter and the nucleotides encoded at positions 21 and 22 are UMP and AMP respectively. When UTP was added, the RNA was elongated to a 21mer (Fig. 2A, lane 2). Addition of either 5-SH-UTP (lane 4) or 5-APAS-UTP (lane 6) also resulted in elongation of the 20mer RNA to a 21mer. The mobilities of the analog-substituted RNAs are somewhat slower than the UMP-containing RNA in this gel system. Control reactions in which no UTP was added (Fig. 2A, lane 1), 100 µM ultrapure ATP was added (not shown) or 1 µM ultrapure ATP, CTP and GTP were added (not shown) resulted in no elongation of the 20mer RNA. Thus, 5-SH-UMP can specifically replace UMP at the 3′-end of the RNA utilizing *E.coli* RNA polymerase. Irradiation of ternary

transcription complexes was carried out for 2 min at 302 nm and the polymerase subunits were then analyzed by SDS–PAGE and autoradiography (Fig. 2B). Only the 5-APAS-UMP-modified RNA resulted in any crosslinking of the RNA polymerase under these conditions, as evidenced by radioactive labeling of the β and β' subunits of the enzyme. These results are consistent with our earlier work. These results suggest that either the mercapto group of 5-SH-UMP did not contact polymerase when at the 3'-end of the RNA or that 5-SH-UTP is not a photoreactive nucleotide analog. Although 4-S-UTP, which has nearly the same absorption maximum as 5-SH-UTP, is photoactivated with this light source (31), we also considered the possibility that 5-SH-UTP might require photoactivation with a different light source, as discussed below.

Analog ionization and transcription by *E.coli* RNA polymerase as a function of pH

The transcription experiments involving the 21mer RNA were carried out at pH 8.0, which is the pH of the transcription buffer we normally use for *E.coli* transcription. At this pH we expected that the thiol group on the analog would be fully ionized, resulting in an analog which carries an additional negative charge as compared with UTP. To determine the pK_a for 5-SH-UTP, the absorption spectra were recorded over a pH range of 5.0–8.0 (Fig. 3A). The absorption at 336 nm is characteristic of the ionized state of the mercapto group (29). At pH 5.0 there was no absorption at 336 nm and therefore the compound was fully protonated and uncharged. Ionization was ~50% at pH 5.6, thus defining the approximate pK_a for 5-SH-UTP.

Although the results shown in Figure 2 indicated that the analog could be incorporated specifically into the 3'-end of the RNA, even when fully ionized, the possibility that the extra negative charge would affect other steps in elongation or termination was investigated. To determine over what pH range this analog could be used with *E.coli* RNA polymerase, the transcription properties of the enzyme were examined from pH 5.0 to 8.0 (Fig. 3B). The template used contained the P1 and P2 promoters of the *E.coli rrnB* operon followed by the strong Rho-independent t_1 terminator, 380 and 260 bp downstream respectively. As expected, transcription was optimal at pH 8.0, but specific transcripts were detected over the entire pH range. The pH range under which the *E.coli* enzyme functions is quite broad and that of other polymerases may be more restrictive. Transcription with 5-SH-UTP was carried out at both pH 6.0 and 8.0 and the RNA pause sites observed with each were the same (not shown).

Because transcription efficiency was greatest at pH 8.0 and it appeared that at least Rho-independent termination at t_1 was unaffected by the presence of the extra negative charge on the analog, the effect of the analog on transcriptional pausing with *E.coli* RNA polymerase was investigated. In addition, the ability of the transcription elongation factor NusA to recognize the transcription complex and enhance pausing was examined (Fig. 4). The pausing assay provides information about the effect of the analog on RNA secondary structure formation, as many pause sites occur downstream from RNA stem–loop structures. For this assay, the T7 A1 promoter was used, from which we have previously determined there is a strong NusA-enhanced pause at position +80, which does occur immediately downstream of an RNA hairpin (13). Synchronously started transcription complexes paused at position +20 were prepared and then the RNA

was elongated in the presence of UTP or 5-SH-UTP, with or without NusA. At the 5-SH-UTP concentration used in this experiment (100 μM) the elongation rate was somewhat slower, but the polymerase paused at the same sites. The major pause at +80 is indicated. The enhancement of pausing by NusA can be seen by the fact that the 80mer RNA persists longer in the presence of NusA (compare the 15 min time points). NusA-enhanced pausing was observed even with complete replacement of UTP by 5-S-UTP in the reaction, indicating that recognition of the complex by this transcription factor was unaffected by the presence of the analog and that 5-SH-UMP could be incorporated at internal positions in the RNA, as well as at the 3'-end. Further, Rho-independent termination at the terminator t_e was normal with the analog, again indicative of normal stem–loop structure formation. The elongation rate with 5-SH-UTP could be increased to that observed with UTP by increasing the concentration of the analog to 1 mM (not shown).

Substrate properties of 5-SH-UTP with T7 RNA polymerase and yeast RNA polymerases I and III

The ability of 5-SH-UTP to function as a substrate for three other RNA polymerases was evaluated. Transcription with bacteriophage T7 RNA polymerase was compared in the presence of 100 μM UTP or 100 μM 5-SH-UTP (Fig. 5A). Results were similar to those obtained with the *E.coli* enzyme. 5-SH-UTP could fully replace UTP as a transcription substrate. Thus 5-SH-UTP can be used to prepare large quantities of any RNA if the gene is cloned into any number of commercially available cloning vectors containing a multiple cloning site downstream from a T7 polymerase promoter.

Utilization of 5-SH-UTP by two eucaryotic enzymes, yeast RNA polymerases I and III, was also tested *in vitro*. Both of the yeast RNA polymerases examined could incorporate 5-SH-UTP into nascent RNA chains (Fig. 5B and C). However, 5-SH-UTP was a rather poor substrate for RNA polymerase I under these conditions (Fig. 5B, lanes 1 and 2). The nuclear yeast RNA polymerases are all quite complex; consisting of ~12–16 subunits (35). Several of the smaller subunits are shared between polymerases, while the largest two subunits are unique to each polymerase. The apparent difference in the extent to which RNA polymerases I and III utilized 5-SH-UTP as a substrate may reflect slight structural differences between catalytic sites of these enzymes.

Post-transcriptional modification of RNA in transcription complexes

To determine whether RNA containing 5-SH-UMP could be specifically modified with a thiol-specific reagent, RNA was synthesized with *E.coli* RNA polymerase with either UTP or 5-SH-UTP as a substrate (Fig. 6). Transcription complexes were then split in half and half of each reaction was treated with APB at pH 7. This reagent reacts fairly specifically with ionized thiol groups. Because the pK_a for 5-SH-UMP is significantly lower than that of cysteine, preferential modification of the RNA might be achieved at pH 7. At this pH, the cysteine thiol groups on proteins should be only slightly ionized. We found that RNA containing UMP or 5-SH-UMP without the addition of APB resulted in only negligible photocrosslinking to the proteins present (Fig. 6, lanes 2 and 6). In the reaction in which the RNA contained UMP and APB was then added, again no crosslinking

of RNA to protein was observed (lane 4). This indicates that either protein thiol groups were not significantly modified by the APB or, if they were modified, none were in contact with any region of the RNA. We cannot distinguish between these two possibilities with this experiment. However, this result supports the conclusion that the protein crosslinking observed when APB was added to 5-SH-UMP-modified RNA was due to specific attachment of the crosslinker to the RNA and not to protein (lane 8).

Evaluation of 5-S-UTP as a photocrosslinking nucleotide analog at different wavelengths

Irradiation at 302 nm of transcription complexes containing 5-SH-UMP at the 3′-end of the 21mer RNA from the T7 A1 promoter did not produce crosslinking of the RNA to the β or β′ subunits of *E.coli* RNA polymerase, as was the case with 21mer RNA containing 5-APAS-UMP at this position (Fig. 2B). Because this could be due to the fact that the thiol group may not have contacted the polymerase at this position in the RNA, we prepared RNA that contained 5-SH-UMP throughout the transcript. Again, no crosslinking to polymerase or Rho was seen when the complexes were irradiated at 302 nm (Fig. 6, lane 6). The absorption maximum for 5-SH-UTP is at 336 nm, with relatively low absorption at 302 nm. However, this is also true for the photoreactive analog 4-S-UTP, and we have found the 302 nm light source to be effective for photoactivation of this compound (31). However, to ensure that the lack of crosslinking with 5-SH-UTP was not caused by irradiation at the wrong wavelength, we compared crosslinking with RNAs containing UMP, 4-S-UMP and 5-SH-UMP. The two light sources used were a broad range lamp with maximum output at 302 nm and a nitrogen laser with monochromatic output at 337 nm. Transcription complexes containing RNA with these different analogs were irradiated for 1 or 10 min and labeling of the β and β′ subunits of the polymerase was examined. As we have seen in the past, both the 302 and 337 nm light sources produced crosslinking to β and β′ with the 4-S-UMP-substituted RNA (Fig. 7). Neither UMP nor 5-SH-UMP-substituted RNA produced any significant crosslinking at either wavelength. A low level of labeling was detected after irradiation of the 5-SH-UMP RNA for 10 min with the 337 nm laser, however, this was <2% of that observed with 4-S-UMP RNA during only 1 min irradiation with the same light source. We therefore conclude that 5-SH-UTP is not itself an effective photocrosslinking nucleotide analog.

DISCUSSION

We have shown that the ribonucleotide analog 5-SH-UTP is an excellent substrate for several RNA polymerases, both procaryotic and eucaryotic. Unlike the photocrosslinking analogs 5-APAS-UTP (13) and 5-APAS-CTP (28), there appears to be no problem with incorporation of two 5-SH-UMPs sequentially in the transcript. With 5-APAS-UTP and 5-APAS-CTP one must include small amounts of UTP or CTP respectively in order to achieve synthesis of full-length RNA transcripts. For example, UTP must be added to at least 5 μM to produce the t_e transcript when 200 μM 5-APAS-UTP is used with 20 μM ATP, CTP and GTP (12). With 5-SH-UTP present at only 100 μM and 20 μM ATP, CTP and GTP, the t_e transcript was produced without addition of any UTP (Fig. 4). Since neither the bis-5-S-UTP (Fig. 1B) nor the ultrapure nucleotides used (not shown) contained any
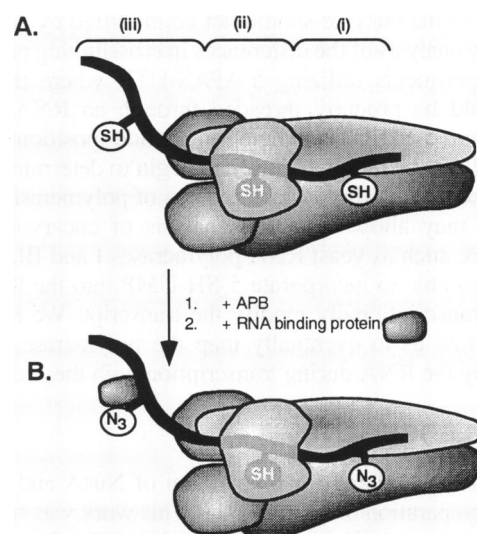

Figure 8. Strategy for determining the contacts of RNA through a transcription complex with 5-SH-UTP and 5-APAS-UTP. RNA is synthesized using 5-SH-UTP as a substrate so that the mercapto groups are incorporated throughout the RNA. The RNA shown can be in three possible environments (**A**). The RNA can be on the surface of the enzyme and accessible to solvent (i), buried within a protein pocket or tunnel and not accessible to solvent (ii) or off the surface of the enzyme and free in the solvent (iii), where other RNA binding proteins can interact. When APB is added to the transcription complexes containing 5-SH-UMP-modified RNA (step 2), only nucleotides in regions (i) and (iii) will react, because the RNA in region (ii) is not accessible to solvent (**B**). Upon irradiation, the modified RNA will crosslink only to binding sites on the surface of the enzyme or to other RNA binding proteins contacting regions on the RNA that have left the enzyme surface.

UTP, production of full-length RNA could only occur by incorporation of analog at adjacent positions. Therefore, using 5-SH-UTP it is now possible to prepare analog-substituted RNA in which the analog can be placed at virtually any position. For preparation of large quantities of analog-tagged RNA, the bacteriophage T7 transcription system can be employed, for which a variety of convenient cloning vectors are commercially available.

Once 5-SH-UMP-substituted RNA has been synthesized, it can then be modified with a variety of thiol-specific reagents (crosslinkers, fluorescent tags, biotin). The RNA can be modified as part of the transcription complex or it can be isolated and then modified in the absence of other protein. We have demonstrated that after modification of 5-SH-UMP-substituted RNA with APB, RNA binding proteins can be crosslinked to the RNA. We observe crosslinking to the transcription termination factor Rho when added to intact transcription complexes (Fig. 6, lane 8). Because the pK_a for 5-SH-UTP is so low compared with protein thiols, it should be possible to develop low pH reaction conditions to preferentially modify the RNA over protein thiol groups. Current models for the structure of the *E.coli* transcription complex include a tight RNA binding region through which it is proposed the RNA is threaded during transcription (17–19). Using the method described here, it should now be possible to carry out experiments to probe for the existence of such a region. If 5-SH-UMP-modified RNA is synthesized with *E.coli* RNA polymerase and then the RNA is modified with APB while still part of the complex, any positions within the RNA that are not on

the surface of the enzyme should not be modified by APB (Fig. 8). Thus, by analysis of the differences in crosslinking patterns in parallel experiments utilizing 5-APAS-UTP, where the cross-linker should be properly threaded through an RNA binding channel, and 5-SH-UTP, where nucleotide positions in the channel will not be modified, one can begin to determine if such a channel exists and, if so, which regions of polymerase form it. 5-SH-UTP may allow structural analysis of eucaryotic RNA polymerases, such as yeast RNA polymerases I and III, as well. It is now possible to incorporate 5-SH-UMP into the RNA and then post-transcriptionally modify the transcript. We hope that this will allow us to eventually map the polymerase subunits contacted by the RNA during transcription with these enzymes.

ACKNOWLEDGEMENTS

REFERENCES

1   Chamberlin,M.J. (1959) *J. Am. Chem. Soc.*, **81**, 4118.
2   Chamberlin,M.J. (1974) *Annu. Rev. Biochem.*, **43**, 721–775.
3   Chamberlin,M.J. (1982) In Boyer,P.D. (ed.), *The Enzymes*. Academic Press, Los Angeles, CA, pp. 61–85.
4   Erie,D.A., Yager,T.D. and von Hippel,P.H. (1992) *Annu. Rev. Biophys. Biomol. Struct.*, **21**, 379–415.
5   Sentenac,A. (1985) In Fasman,G.D. (ed.), *Critical Reviews in Biochemistry*. CRC Press, Boca Raton, FL, pp. 31–90.
6   von Hippel,P.H., Bear,D.G. and Morgan,W.D. (1984) *Annu. Rev. Biochem.*, **53**, 389–446.
7   Ishihama,A. (1993) *J. Bacteriol.*, **175**, 2483–2489.
8   Friedman,D.I. (1992) *Curr. Opin. Genet. Dev.*, **2**, 727–738.
9   Greenblatt,J., Nodwell,J.R. and Mason,S.W. (1993) *Nature*, **364**, 401–406.
10   Das,A. (1993) *Annu. Rev. Biochem.*, **62**, 893–930.
11   Bowser,C. and Hanna,M.M. (1991) *J. Mol. Biol.*, **220**, 227–239.
12   Dissinger,S. and Hanna,M.M. (1990) *J. Biol. Chem.*, **265**, 7662–7668.
13   Dissinger,S. and Hanna,M.M. (1991) *J. Mol. Biol.*, **219**, 11–25.
14   Hanna,M.M. (1993) *Cell. Mol. Biol. Res.*, **39**, 393–399.
15   Liu,K. and Hanna,M.M. (1995) *J. Mol. Biol.*, **247** in press.
16   Zhang,Y. and Hanna,M. (1994) *J. Bacteriol.*, **176**, 1787–1789.
17   Chan,C.L. and Landick,R. (1994) In Conaway,R.C. and Conaway,J.W. (eds), *Transcription: Mechanisms and Regulation*. Raven Press, New York, NY, pp. 297–321.
18   Chamberlin,M.J. (1994) In *The Harvey Lectures*. Wiley-Liss, New York, NY, pp. 1–21.
19   Nudler,E., Goldfarb,A. and Kashlev,M. (1994) *Science*, **265**, 793–796.
20   Hanna,M.M., Dissinger,S., Williams,B.D. and Colston,J.E. (1989) *Biochemistry*, **28**, 5814–5820.
21   Hanna,M.M. (1989) *Methods Enzymol.*, **180**, 383–409.
22   Burgess,R.R. and Jendrisak,J.J. (1975) *Biochemistry*, **14**, 4634–4638.
23   Gonzalez,N., Wiggs,J. and Chamberlin,M.J. (1977) *Arch. Biochem. Biophys.*, **182**, 404–408.
24   Thayer,G.C. and Brosius,J. (1985) *Mol. Gen. Genet.*, **199**, 55–58.
25   King,K., Benkovic,S. and Modrich,P. (1989) *J. Biol. Chem.*, **254**, 11807–11815.
26   Wright,D.J., King,K. and Modrich,P. (1989) *J. Biol. Chem.*, **264**, 11816–11821.
27   Zhang,Y. and Hanna,M.M. (1995) *Protein Expr. Purif.*, submitted for publication.
28   Hanna,M.M., Zhang,Y., Reidling,J.C., Thomas,M.J. and Jou,J.(1993) *Nucleic Acids Res.*, **21**, 2073–2079.
29   Bardos,T.J. and Kalman,T.I. (1966) *J. Pharm. Sci.*, **55**, 606–609.
30   Burns,J.A., Butler,J.C., Moran,J. and Whitesides,G.M. (1991) *J. Org. Chem.*, **56**, 2648–2650.
31   Tanner,N.K., Hanna,M.M. and Abelson,J. (1988) *Biochemistry*, **27**, 8852–8861.
32   Riggs,D.L., Peterson,C.L., Wickham,J.Q., Miller,L.M., Clarke,E.M., Crowell,J.A. and Sergere,J.-C. (1995) *J. Biol. Chem.*, **270** in press.
33   Pavco,P.A. and Steege,D.A. (1990) *J. Biol. Chem.*, **265**, 9960–9969.
34   Schagger,H. and Jagow,G.v. (1987) *Anal. Biochem.*, **166**, 368–379.
35   Thuriaux,P. and Sentenac,A. (1992) In Jones,E.W. Pringle,J.R. and Broach,J.R. (eds), *Gene Expression*. Cold Spring Harbor Press, Plainview, NY, pp. 1–48.