Sun, J. S., Asseline, U., Rouzaud, D., Montenay-Garestier, T., Thuong, N. T., & Hélène, C. (1987) *Nucleic Acids Res. 15*, 6149-6158.
Thuong, N. T., & Chassignol, M. (1987) *Tetrahedron Lett. 28*, 4157-4161.
Thuong, N. T., Chassignol, M., Lancelot, G., Mayer, R., Hartmann, B., Leng, M., & Hélène, C. (1981) *Biochimie 63*, 775-784.
Thuong, N. T., Asseline, U., Roig, V., Takasugi, M., & Hélène, C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5129-5133.
Toulmé, J. J., Krisch, H. M., Loreau, N., Thuong, N. T., & Hélène, C. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1227-1231.

Photoaffinity Labeling of *Escherichia coli* RNA Polymerase/Poly[d(A-T)] Transcription Complexes by Nascent RNA†

Thomas M. Stackhouse and Claude F. Meares*
Chemistry Department, University of California, Davis, California 95616
Received September 14, 1987; Revised Manuscript Received December 22, 1987

ABSTRACT: To elucidate the molecular interactions during transcription by *Escherichia coli* RNA polymerase, we have performed a quantitative analysis of the photoaffinity labeling produced by an aryl azide positioned at the leading (5′) end of the nascent RNA. Macromolecular contacts on the path of RNA across the transcription complex containing the template poly[d(A-T)] are observed as a function of the length of the transcript. Quantitative analysis provides the percent yield of photoaffinity labeling in the transcription complex by each length of RNA. Significant yields are observed for DNA, the β/β' subunits (analyzed together), and the σ subunit. The α subunit is not labeled under these experimental conditions. The DNA template is labeled by the leading ends of RNA molecules 5-18 bases long, with yields ranging from 1% to 6%. Photoaffinity labeling of poly[d(A-T)] is also observed for many transcript lengths longer than 18 nucleotides, but the yields are too low to quantitate. Labeling of the β/β' subunits occurs with ≈50% yields for transcripts of lengths ≥12 nucleotides; low but significant labeling yields of 1-8% by shorter RNAs (3-10 nucleotides) are observed. Labeling of the σ subunit is detectable for transcripts from 7 to more than 19 nucleotides long; quantitative measurements were possible up to the 19-mer. The RNAs most likely to be photoattached to the σ subunit are 9-12 nucleotides long, with a maximum photoaffinity labeling yield of 15% by the decanucleotide. These results modify the conclusions of previous work concerning the release of σ from an *E. coli* RNA polymerase/poly[d(A-T)] transcription complex [Hansen, U. M., & McClure, W. R. (1980) *J. Biol. Chem. 255*, 9564-9570]. The photoaffinity labeling of σ in poly[d(A-T)] transcription complexes differs from the results observed with DNA containing either the λ P_R or the T7 A1 promoter [Bernhard, S. L., & Meares, C. F. (1986) *Biochemistry 25*, 5914-5919], providing further evidence that the interaction between the nucleic acids and the σ subunit in the transcription complex depends on the nucleotide sequence.

The control of gene expression at the level of transcription has been studied most extensively with *Escherichia coli* RNA polymerase. This oligomeric enzyme (EC 2.7.7.6) catalyzes the synthesis of ribonucleic acid from a deoxyribonucleic acid template. RNA polymerase from *E. coli* contains five major subunits, with a total molecular weight of 449 000. The primary structures of all the subunits have been determined: α (M_r 36 512; Ovchinnikov et al., 1977); β (M_r 150 619; Ovchinnikov et al., 1981); β' (M_r 155 162; Ovchinnikov et al., 1982); and σ (M_r 70 263; Burton et al., 1981). The core enzyme contains four subunits ($\alpha_2\beta\beta'$) and is capable of elongating, but not efficiently initiating, RNA transcripts from promoter sites on DNA. Efficient initiation of a transcript at promter sites on DNA requires the holoenzyme, which contains the core RNA polymerase and the σ subunit. The presence of another subunit (ω) has also been observed [see Gentry and Burgess (1986) and references cited therein]; the function of ω is not yet established.

It was discovered by Travers et al. (1969) that σ can be released shortly after the initiation of transcription and subsequently σ can bind to another core enzyme to initiate another transcript. This is referred to as the σ cycle (Lewin, 1983). A detailed understanding of the mechanism of the σ cycle and its relation to the initiation reaction has interested many investigators (Chamberlin, 1974; Hansen & McClure, 1980; Shimamoto et al., 1986). It is believed that release of the σ subunit is accompanied by the formation of a stable elongation complex containing core enzyme, DNA template, and nascent RNA. The elongation complex could subsequently bind other factors involved in control of transcription and translation (McClure, 1985; Greenblatt et al., 1987). For example, it has been shown that nusA protein can bind to the elongation complex only after σ release (Greenblatt & Li, 1981).

In order to determine the RNA length at which σ is released, Hansen and McClure (1980) performed an elegant experiment measuring the amount of σ present in a mixture of *E. coli* RNA polymerase/poly[d(A-T)] transcription complexes containing various transcript lengths. By careful comparison of the amount of σ released with the lengths of RNA present, these authors concluded that σ is released quantitatively from the complex by the time a transcript of eight or nine nucleotides has been produced on poly[d(A-T)].

Other investigators have examined the release of σ from transcription complexes that contain the A1 promoter of bacteriophage T7 (Shimamoto et al., 1986). This group an-

†Supported by Research Grant GM 25909 from the National Institute of General Medical Sciences, NIH.

0006-2960/88/0427-3038$01.50/0

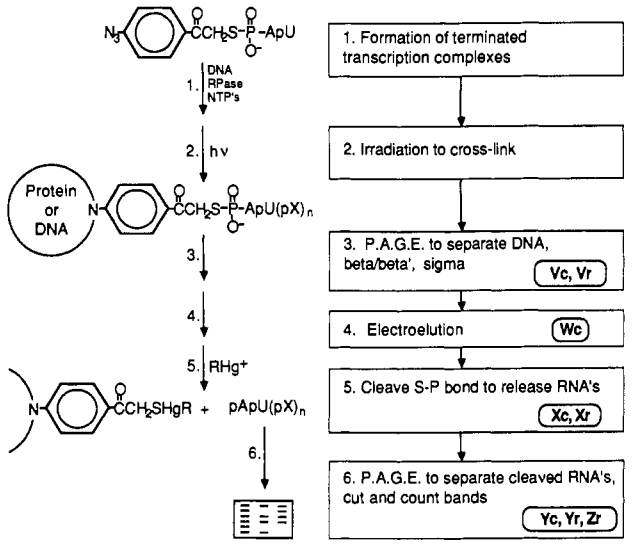

FIGURE 1: Experimental outline, showing the structure of the photoaffinity probe 5'-[[(4-azidophenacyl)thio]phosphoryl]adenylyl(3'–5')uridine at the top. The points where yield factors (V_c, V_r, W_c, etc.) accompanying each process were calculated are indicated. These were necessary for the percent yield determinations, as described under Methods.

alyzed the release of σ as a function of time and proposed a two-step model that involves a fast triggering step (<1 s) followed by a slower dissociation (mean lifetime ≈5 s).

Recently, our laboratory has used an aryl azide photoaffinity probe at the leading (5') end of the RNA (Figure 1) to qualitatively follow the path of the transcript as it moves through an *E. coli* transcription complex containing T7 A1 or λ P_R DNA (Bernhard & Meares, 1986a; Hanna & Meares, 1983a,b). The photoaffinity probe is a dinucleotide, allowing specific initiation of transcription from promoters whose transcripts begin with A-U. The aryl azide is converted to a highly reactive nitrene by irradiation with UV light [wavelength (λ) >300 nm]. This nitrene can insert nonselectively into many types of chemical bonds, thus providing information about which molecules are within its immediate vicinity (Bayley & Staros, 1984; Bayley & Knowles, 1977; Knowles, 1971). The probe also contains a sulfur–phosphorus bond that is cleaved by mercurials. Therefore, the enzyme subunits, once cross-linked to the nascent RNA, can be isolated and the RNA cleaved from the protein to determine the RNA chains whose 5' ends were adjacent to each subunit.

The information obtained from the A1 promoter of bacteriophage T7 ΔD111 and ΔD123 DNA (Hanna & Meares, 1983a,b) indicates the following: when the RNA is 3–12 bases long, the DNA template is heavily photoaffinity labeled; when the trinucleotide transcript is made in large quantities, the σ and β subunits are just detectably labeled, but not β' or α; for RNA lengths of 4–12 nucleotides, β and β' are labeled but not σ or α. For RNA lengths 12 nucleotides or longer (and as long as 94 nucleotides on T7 ΔD123), the β and β' subunits are heavily labeled, but the other macromolecules in the transcription complex are not labeled significantly. thus, σ appears to be in contact with the 5' end of the RNA trinucleotide made from the T7 A1 template, but not with longer RNAs.

Further studies by Bernhard and Meares (1986a) have shown that the photoaffinity labeling of the σ subunit depends on the sequences of the nucleic acids involved. These authors compared the T7 A1 and λ P_R transcription complexes with the finding that on the λ P_R template the 5' ends of RNAs with lengths of 9–13 nucleotides heavily labeled σ. Photoaffinity labeling results for β, β', and α were similar on both templates, suggesting that contacts between RNA and the subunits of the core enzyme are not strongly dependent on the nature of the DNA.

Probes with lengths of 9–13 nucleotides label β, β', and DNA on both templates and thus appear to be free to label any molecule in their immediate surroundings. Therefore, these results strongly suggest that the reason σ is not labeled by these lengths on T7 A1 is that σ dissociates from the T7 A1 transcription complex much earlier than from λ P_R. The interaction responsible for σ dissociation on the T7 A1 template evidently occurs before the transcript is nine bases long, while the 5' end of the RNA is still hybridized to DNA (Hanna & Meares, 1983b). Therefore, it appears to us that the sequence of the RNA/DNA hybrid is likely to play an important role.

On the basis of current knowledge of the mechanism of protein/nucleic acid interactions (Seeman et al., 1976; von Hippel, 1979, 1982), the key interactions are expected to involve formation of specific hydrogen bonds between amino acid side chains and hydrogen-bond donors and acceptors on the nucleic acid. Bernhard and Meares proposed a hypothesis that depends on the spatial distribution of bridging hydrogen-bond donors and acceptors in the major groove of the RNA/DNA hybrid in the transcription complex. Examination of their sequences shows that a major difference between the T7 A1 and λ P_R templates is an alternating hydrogen-bond donor/acceptor pattern in the early transcript region of T7 A1 that is not present in λ P_R. The bridging hydrogen-bond donor/acceptor pattern in the early transcript region of T7 A1 just matches that in poly[d(A-T)]; from the experiments of Hansen and McClure (1980) regarding σ release from poly[d(A-T)], we would expect photoaffinity labeling of σ not to occur for RNA lengths longer than eight or nine nucleotides.

Here, we discuss the initial results of our endeavor to determine more precisely the nucleic acid sequence dependence of photoaffinity labeling with this 5' RNA probe in *E. coli* transcription complexes. In this context, it is important to measure the percent yields of photoaffinity labeling on the subunits in transcription complexes as a function of RNA chain length. This adds significantly to the information obtained from these experiments, because it corrects for inequalities in amounts of different RNA chain lengths present, for various losses during sample treatment, and for the limited dynamic range of film. This approach will allow quantitative comparison of photoaffinity labeling using different DNA templates in the future.

Experimental Procedures

Materials

All reagents and solvents were the purest available and used without further purification unless otherwise noted. Nanopure water (Barnstead) was used throughout. *E. coli* MRE 600 cells were purchased from Grain Processing Corp. Bactotryptone and yeast extract were from Difco. High-performance liquid chromatography (HPLC)[1]-purified ribonucleoside triphosphates were purchased from ICN. Ultrapure urea and acrylamide were purchased from Schwarz/Mann. Soluble RNA, phenylmercuric acetate, dithiothreitol, tetramethylethylenediamine, methylenebis(acrylamide), cordycepin triphosphate cordycepin triphosphate (3'-deoxy-ATP, a chain-

[1] Abbreviations: Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; HPLC, high-performance liquid chromatography; $NaDodSO_4$, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis.

terminating ATP analogue), agar, Tris, ethidium bromide, and agarose were purchased from Sigma. *p*-Azidophenacyl bromide was from Pierce, adenosine 5′-*O*-(3-thiotriphosphate) was from Boehringer-Mannheim, and 3′-*O*-methyl-UTP (RNA chain terminator) was from P-L Biochemicals. [α-^{32}P]ATP was purchased from Amersham. Kodak XAR films were used without intensifying screens for autoradiography.

Buffers. These were as follows: *Buffer A* contained 80 mM Tris-HCl, pH 7.9, 5 mM 2-mercaptoethanol, 0.1 mM Na_2EDTA, and 50% (v/v) glycerol; *buffer B* contained 10 mM Tris-HCl, pH 8.0, and 1 mM Na_2EDTA; *buffer C* contained 50 mM Tris-HCl, pH 7.9, 5 mM 2-mercaptoethanol, 10 mM NaCl, 10 mM $MgCl_2$, and 5% (v/v) glycerol; *buffer D* contained 0.5 mL of buffer F, 2.9 mL of 9–10 M urea, 0.5 g of sucrose, 12 mg of $NaDodSO_4$, 0.1 mL of 0.1% (w/v) bromphenol blue, and 10 mM dithiothreitol (added just prior to use); *buffer E* contained, in 1 L, 4.4 mL of ethanolamine, 4.5 g of glycine, and 1.0 g of $NaDodSO_4$ (pH 9.7); *buffer F* contained 18.6 mL of triethanolamine, 8 mL of concentrated HCl, and 96 g of urea in 200-mL total volume, pH 7.5; *buffer G* contained 89 mM Tris–borate (pH 8.3), 1 mM Na_2EDTA, and 7 M urea; *buffer H* contained 25 mM Tris-HCl, pH 8.5, 10 mM glycine, and 0.1% $NaDodSO_4$.

Cleavage Solution. First, 0.1% (w/v) $NaDodSO_4$ is saturated with phenylmercuric acetate by shaking with an excess of the pure solid for a minimum of 4 h at room temperature in the dark. The mixture is briefly centrifuged to pellet the excess solid, and the precipitate is discarded. Soluble RNA (1 mg/mL) is added to the cleavage solution as a carrier. The solution is used within 1 day after preparation.

Methods

RNA Polymerase. The enzyme was purified from *E. coli* MRE 600 cells according to the method of Burgess and Jendrisak (1975) as modified by Lowe et al. (1979). It was dialyzed against buffer A and stored at −20 °C. Protein concentration was determined by the method of Burgess (1976).

DNA Template. Poly[d(A-T)] was purchased from Sigma and was extracted with phenol/chloroform, precipitated with ethanol, resuspended in buffer B, and stored at −20 °C. The concentration was determined by measuring the absorbance at 260 nm (one A_{260} unit = 50 μg/mL) and using an average molecular weight of 7.5×10^5.

Dinucleotide Photoaffinity Probe. 5′-[[(4-Azidophenacyl)thio]phosphoryl]adenylyl(3′–5′)uridine (N_3RSpApU) (Figure 1) was synthesized as reported by Hanna and Meares (1983a). Briefly, a reaction was used which took advantage of the abortive initiation reaction of RNA polymerase using a poly[d(A-T)] template and using UTP and adenosine 5′-*O*-thiophosphate as substrates. This allowed the formation of the dinucleotide 5′-*O*-thiophosphoryladenylyl(3′-5′)uridine, which was isolated by HPLC. The dinucleotide was then alkylated by *p*-azidophenacyl bromide in reduced light. The excess *p*-azidophenacyl bromide was extracted with isobutyl alcohol, followed by ethyl ether. The protoprobe was isolated by HPLC, and the product concentration was determined by using $\epsilon_{300} = 2 \times 10^4$ M^{-1} cm^{-1}.

Preparation of Transcription Reactions. The experimental procedure is outlined in Figure 1. All transcription reactions were carried out in reduced light. Two separate transcription reactions were performed, one to terminate the transcript at adenylate residues and the second to terminate the transcript at uridylate residues. The main goal was to produce an evenly distributed set of RNA lengths in the transcript mix so that all are represented for photoaffinity labeling. First, two master mixes were prepared: a preinitiation mix and an initiation mix.

The *preinitiation mix* was prepared by combining, in a total volume of 14.0 μL, 1.4 μM RNA polymerase, 0.36 μM poly[d(A-T)] (based on M_r 7.5×10^5), and 470 μM N_3RSpApU in buffer C. This was incubated at 37 °C for 10 min to allow the formation of a ternary complex.

The *initiation mix* was prepared from unlabeled UTP and [α-^{32}P]ATP (3000 Ci/mmol) at concentrations of 5 and 3 μM, respectively, in a total volume of 12 μL. The preinitiation mix was mixed with the initiation mix and incubated at 37 °C for 15 s to allow short transcripts to be produced that contain the radioactive label. The resulting solution is called the *pulsed mix*.

A 12-μL aliquot from the pulsed mix was added to elongation mix A, forming a complete reaction mixture (48 μL) which contained 18 μM ATP, 18 μM UTP, and 500 μM cordycepin triphosphate, in buffer C. Another 12-μL aliquot from the pulsed mix was added to elongation mix U, forming a complete reaction mixture (48 μL) which contained 18 μM ATP, 18 μM UTP, and 500 μM 3′-*O*-methyl-UTP, in buffer C. The 4-fold dilution of the pulsed mix and 18-fold increase in cold ATP concentration served to limit the increase of specific activity with RNA chain length during elongation. These complete reaction mixtures were incubated for 10 min at 37 °C. At the end of this elongation period, 4-μL aliquots were removed from each reaction mixture and added to 5 μL of RNA sample buffer [5× buffer G containing the tracking dyes bromphenol blue and xylene cyanol (0.25% w/v)] and 9 μL of cleavage solution. These unirradiated RNA samples were incubated in the dark for 12 h to cleave the probe. Subsequently, 2 μL of 1 M dithiothreitol was added to these RNA samples, which were incubated for 1 h in the dark at room temperature to reduce the azide, and run on an RNA sequencing gel to confirm the lengths of RNA produced.

Unirradiated control protein samples were prepared by removing 18 μL from each final transcript mix and adding it to an equal volume of 2× buffer D. These protein samples were allowed to incubate at room temperature for 1 h in the dark before electrophoresis on a protein gel.

After the control samples had been taken, the A and U reactions were transferred to a borosilicate tube and irradiated for 30 s in a Rayonet photochemical reactor ($\lambda > 300$ nm). Following irradiation, samples were again removed for both RNA sequencing and protein analsyis, as was done for the unirradiated controls above.

Protein Gels. The protein gels used in these experiments employed the $NaDodSO_4$–urea system of Wu and Bruening (1971). The running gel was 6% acrylamide with a length of 20 cm. Except for β and β', which moved as a single broad band, this was found to be sufficient to separate the polymerase subunits from each other, the DNA, and the small transcripts. Gels were run for 1 h at 80 V and then at 30 mA until the bromphenol blue tracking dye had run off the bottom of the gel (approximately 3 h). After electrophoresis was complete, the gels were wrapped in plastic wrap and autoradiographed at room temperature for 1–2 h with Kodak XAR film.

Electroelution and Cleavage. The photoaffinity-labeled subunits and DNA, which were located by autoradiography and comparison to stained marker lanes, were excised from the protein gel. The individual gel pieces were transferred to 1.5-mL polypropylene Eppendorf centrifuge vials and counted (Cerenkov radiation) in a Beckman LS 6800 scintillation counter prior to electroelution.

Electroelution was carried out by placing the individual gel pieces plus 200 μL of buffer H into dialysis tubing (molecular

weight cutoff 12 000–14 000). The dialysis bags were placed into a Bio-Rad Trans Blot apparatus eluted at 10 °C and 75 V (1 mA) for 8 h. The eluted protein–RNA conjugates were transferred to 0.5-mL Eppendorf vials. Each dialysis bag was rinsed with 200 μL of freshly prepared cleavage solution. The rinse was pooled with the original eluent; these solutions were counted as were the remaining dialysis bags and gel pieces for each sample. These counts were then used to determine the efficiency of the elution process. Cleavage was allowed to proceed for 18–24 h at room temperature in the dark.

Following the elution and cleavage, the samples were lyophilized to dryness (8 h, 20–50 mTorr) and then resuspended in 10 μL of H_2O and 10 μL of RNA sample buffer (above) and counted. The samples were then briefly centrifuged to settle the precipitate, and the supernatant was removed and counted. These counts were then used to determine the efficiency of the cleavage reaction. Six microliters of each sample was then run on an RNA sequencing gel to determine the lengths of transcript that had cross-linked to the protein or DNA.

RNA Sequencing Gels. Total RNA samples and RNA cleaved from protein or DNA were analyzed on 40 cm × 0.75 mm 25% acrylamide gels [1:29 methylenebis(acrylamide): acrylamide ratio] containing buffer G (Carpousis & Gralla, 1980). Gels were allowed to polymerize for 2 h and were then preelectrophoresed at 1000 V for a minimum of 4 h before use. For the actual analysis of the samples, each gel was run at 800–1000 V until the bromphenol blue had run 27.5 cm from the origin. After electrophoresis, the gels were placed between two sheets of plastic wrap and autoradiographed at −80 °C witout an intensifying screen.

Quantitation of Transcripts. The individual lanes from the RNA sequencing gels were cut out and sliced, starting just below the trinucleotide and working toward the top of the gel; each slice was 3 mm wide. The individual gel slices from each lane were counted (Cerenkov radiation). The gel slices were placed flat on the bottom of the scintillation vials to prevent any differences in counting efficiency due to geometric factors. The gel slices typically extended up the lane to transcripts 20 nucleotides long before resolution became uncertain. Total counting time for a complete gel (eight lanes) was about 48 h. This rather laborious manual procedure may be avoided if automated instruments are available.

The percent yield of photoaffinity labeling can be calculated by comparing the counts in the lanes that contained RNA released from protein or DNA to the counts in the marker lanes, which were accurate dilutions of an irradiated and cleaved RNA sample from the original transcription mix:

$$\% \text{ yield} = 100F\frac{\text{cpm of cleaved transcript}}{\text{cpm of marker transcript}}$$

where

$$F = V_r X_r Y_r Z_r / V_c W_c X_c Y_c$$

and V_c and V_r = volume removed from the original transcription mix for subunit analysis or RNA marker, respectively, W_c = electroelution efficiency, X_c and X_r = cleavage reaction efficiency for subunit samples and RNA markers, respectively, Y_c and Y_r = fraction of final sample loaded onto the RNA sequencing gel for subunit sample and RNA marker, respectively, and Z_r = dilution factor for RNA markers.

The points at which these factors were measured are shown in Figure 1. In this way, the number of moles of a particular transcript in the transcription mix could be directly related to the number of moles of that same transcript photo-cross-linked to enzyme subunits or DNA.

The percent photolabeling yield may be analyzed as a ratio of pseudo-first-order chemical reaction rate constants. For example, the yield for attachment of the nitrene at the leading end of RNA to the σ subunit is

$$\% \text{ yield}_\sigma = \frac{k_\sigma}{k_\sigma + k_{\beta\beta'} + k_{\text{DNA}} + k_{\text{other}}}$$

The rate constant k_σ (for example) is proportional to the effective concentration of σ in the vicinity of the nitrene. This effective concentration depends on the rate of collisions between the nitrene and amino acid residues on σ, and also on the reactivities which those residues exhibit toward the nitrene (Bayley & Staros, 1984). Since the residues involved have not yet been identified, we cannot separate the collision frequencies from the intrinsic reactivities at this time. However, it is clear that the intrinsic reactivities of the competing residues do not play a dominant role in the results, because in most instances the nitrenes on oligonucleotides of a given length (e.g., the decamer) label as many as three or four different macromolecules (β or β′, σ, DNA). Thus, it is reasonable to expect that the photoaffinity labeling percent yields at least qualitatively reflect the frequency of encounters between the nitrene and the labeled macromolecule.

A related question is whether the composite rate constant for all reactions that do not lead to cross-linking, denoted k_{other} in the equation above, is likely to vary with chain length. The experiments of Bernhard and Meares (1986b) on the accessibility of the probe to solvent molecules answer this question affirmatively. Furthermore, the formation of secondary structures (e.g., hairpin loops) in the RNA is likely to vary with chain length; this will cause the self-labeling of RNA to change accordingly. Thus, for probes on RNAs of different lengths in transcription complexes, it is to be expected that the total fraction of nitrene reacting with RNA or solvent molecules (H_2O, Tris, mercaptoethanol, etc.) can vary widely.

Results

Figure 2 shows a typical autoradiogram observed for the protein gel. It is clearly evident that σ, β and β′, and DNA were cross-linked to the radioactive message; however, no labeling is observed on the α subunit. This is analogous to the labeling pattern seen on the λ P_R template (Bernhard & Meares, 1986a). The β and β′ subunits were analyzed together; for T7 A1, it was found that upon isolation of the two subunits, both were labeled with the same RNA lengths except for the trinucleotide (Hanna & Meares, 1983b). The control experiments in Figure 2 show that the protein is not labeled without irradiation. A small amount of radioactivity appears in the DNA in the unirradiated control experiment due to the ability of RNA polymerase to add nucleotides to the 3′ ends of DNA, including any nicks along the DNA. This labeling is not susceptible to cleavage by the mercurial that will follow, so it does not interfere with these studies (see Experimental Procedures). For the gel shown in Figure 2, the labeled bands in the A lane contained 390 526 cpm for β, β′, 71 528 cpm for σ, and 54 410 cpm for the DNA. The U lane contained 714 070 cpm for β, β′, 150 530 cpm for σ, and 90 457 cpm for the DNA. These numbers are cited merely to indicate that sufficient radioactivity was present for accurate counting statistics.

An autoradiogram of an RNA sequencing gel is presented in Figure 3. This gel shows the transcripts whose 5′ ends were photoattached to the various parts of the transcription complex. Labeling of σ by transcripts 7–12 bases long is readily ob-

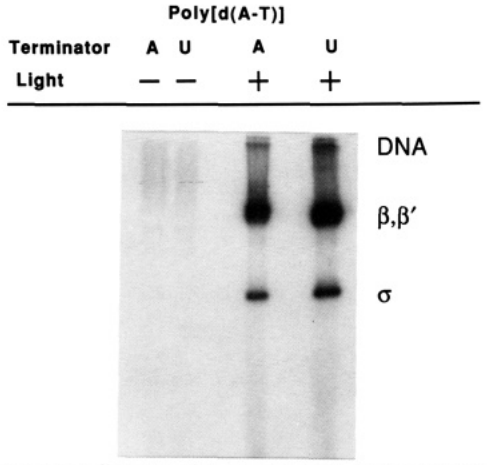

FIGURE 2: Autoradiogram of the $NaDodSO_4$–urea–polyacrylamide gel indicating the separation and photoaffinity labeling of the transcription complex subunits. The first two lanes are the unirradiated controls, and the last two lanes are the photolyzed reactions, each with the RNA chain terminated at an A or U as indicated.

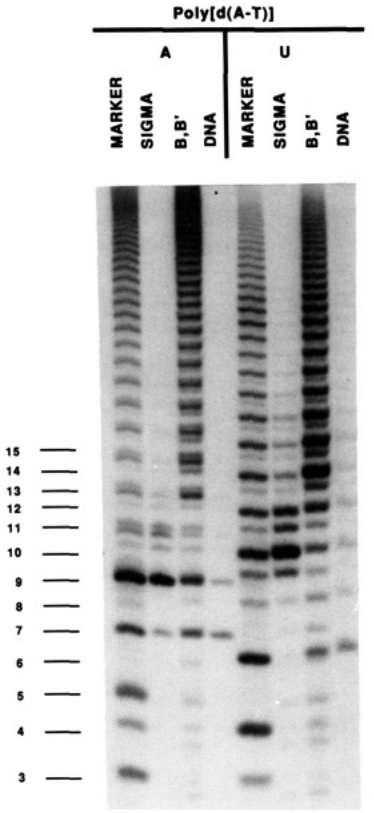

FIGURE 3: Autoradiogram of the 25% polyacrylamide RNA sequencing gel indicating all the transcript lengths present in the original transcription mix ("Marker") and the RNA cleaved from the individual subunits of the complex ("SIGMA"; "B,B′"; and "DNA"). As in Figure 2, the presence of the specific terminators is indicated by A or U. Note that transcripts other than those expected for specific termination are observed. These are due to abortive initiation or to the low concentrations of terminator, ATP, and UTP present; they were not used in the subsequent quantitative analysis (Figure 4).

servable, and a small but significant amount of labeling arising from longer lengths of RNA can also be seen. Adequate radioactivity was available for good counting statistics; for example, the band containing the RNA 10-mer cleaved from σ contained 2452 cpm. β and β' were analyzed together, due to lack of separation on the protein gel; they are labeled by nearly all RNA chain lengths, with various efficiencies. These results for β and β' are similar to the qualitative labeling patterns seen on the λ P_R and T7 A1 templates mentioned

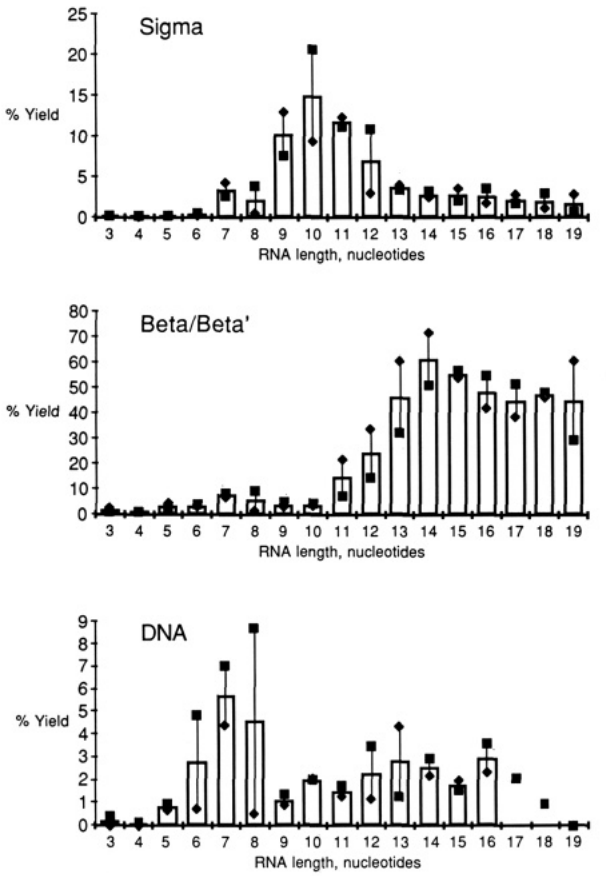

FIGURE 4: Percent photoaffinity labeling yield of σ, β/β', or DNA, analyzed from poly[d(A-T)] transcription complexes as a function of transcript length. The results of two independent experiments are shown (diamonds, squares) along with their average (columns). Note that the vertical scales are quite different in the three panels. For DNA, the percent yields for the 17- and 18-mers could only be determined in one experiment.

above. The DNA sample shows a low level of labeling by transcript lengths from 6 to >20 nucleotides. The labeling of σ and DNA by transcript lengths >13 nucleotides has not been observed on other DNAs. A possible explanation for why it occurs on poly[d(A-T)] is given below.

Figure 4 shows the photoaffinity labeling yields for each of the macromolecules analyzed, as a function of transcript length. This provides a quantitative analysis of what is seen on the autoradiogram of the RNA sequencing gel in Figure 3. The results of two independent experiments are shown in Figure 4, together with their average. The data in Figure 4 appear somewhat scattered, principally because of the limitations on resolution imposed by the 3-mm slice width when comparing separate gels. The σ subunit is labeled most heavily by transcript lengths of 9 through 12 nucleotides, with a maximum yield of approximately 15% at a transcript length of 10 nucleotides. The labeling yield for β and β' is 1–8% for transcripts 3–10 nucleotides long and then dramatically increases to approximately 60% at a transcript length of 14 nucleotides. The percent yield on β and β' remains high for the rest of the lengths analyzed. The DNA labeling fluctuates around 1–6% for the early transcripts of 5 through 18 nucleotides and becomes too low to quantitate for longer RNAs.

DISCUSSION

σ Release. It has been established that the σ subunit of RNA polymerase can be released from the transcription complex after initiation of RNA synthesis (Travers & Burgess, 1969). Hansen and McClure (1980) studied the length of

transcript at which the σ subunit is released from a ternary complex using poly[d(A-T)] as the template. Studies of this type are particularly difficult, because RNA polymerase cannot be induced to pause uniquely at each nucleotide position in an RNA chain. Rather, each DNA template posseses some discrete pause sites that can be conveniently studied but are not necessarily located near the point of σ release. After using the chain terminator 3′-deoxy-ATP to produce a mixture of chain lengths and determining the amounts of the various lengths of RNA produced and the amounts of dissociated σ, Hansen and McClure developed a mathematical model of the process. This model indicated that all of the σ subunits are released at a transcript length of eight or nine nucleotides.

However, the data in Figures 3 and 4 clearly indicate that σ is still present in at least some transcription complexes when the transcript is 10, 11, or 12 nucleotides long. σ is photoaffinity labeled by about 15% of the RNA chains 10 nucleotides long; this implies that at least 15% of the transcription complexes still contain σ at that transcript length.

Methodology. Bayley and Staros (1984) point out that the extent of labeling in different photoaffinity experiments varies from almost 100% to less than 1%. The accessibility of the bound photoprobe to solvent molecules, which can compete with the macromolecule for reaction with the nitrene, may lead to low photoaffinity labeling yields. Also, the covalent bond formed between the photoprobe and the macromolecule may not always be sufficiently stable to permit isolation of a cross-linked product. A third point specifically concerning RNA polymerase is the abortive initiation reaction, which leads to repeated production and dissociation of oligonucleotide transcripts, in some cases up to 12–13 nucleotides in length (Carpousis & Gralla, 1980; Levin et al., 1987). Having been released from the transcription complex, these abortive products react with the solvent; since the percent yield calculation counts all the RNA molecules produced, this contributes to the low percent yields observed for short RNA.

The reason for the discrepancy as to the exact transcript length at which σ is released could be due to the different experimental procedures used. Hansen and McClure performed experiments that involved separation of σ from ternary enzyme/DNA/RNA complexes using gel filtration through Bio-Gel A-1.5m. If the σ-containing transcription complexes were only marginally stable, the added manipulations might lead to the artificial release of σ. Photoaffinity cross-linking is performed in homogeneous solution, with no prior chromatographic separation of components. Second, the Hansen and McClure analysis of the RNA products involved quantitation of the transcripts by autoradiography and subsequent scanning of the autoradiogram, after separation of the different lengths by 20% acrylamide/urea–PAGE. Recently, it has been observed that 25% gels afford more reliable analysis of short oligonucleotides (Levin et al., 1987). In this investigation, we used the RNA sequencing 25% PAGE of Carpousis and Gralla (1980) to analyze the transcript lengths and quantitated the transcripts by cutting and counting individual gel slices, compared to a marker of total RNA from the same reaction.

It might be argued that the presence of the photoaffinity probe at the 5′ end of the RNA somehow perturbs the system such that some of the σ remains bound. While it is difficult to rule this out entirely, we believe that the high sensitivity of these experiments to subtle variables such as the DNA base sequence supports the notion that the properties of the photoprobe do not interfere with the results.

Comparison of Figures 3 and 4 shows the strengths and weaknesses of autoradiography versus counting. Autoradiography detects small amounts of radioactivity with high sensitivity and resolution, such as the RNA chains >15 nucleotides long photoattached to σ and to DNA in the U-terminated lanes of Figure 3. Also, autoradiography reveals the presence of other RNA bands caused by abortive initiation and low substrate concentration, but counting (Figure 4) makes the actual quantitative relationships between different bands clearer and thus establishes their relative significance. For example, it is evident that the RNA 9-mer labels β/β' more heavily than the 7-mer in the A-terminated lanes of Figure 3. However, more 9-mer than 7-mer was present in the reaction mixture, and the actual percent yields show that the 7-mer labels β/β' more efficiently (8%) than the 9-mer (4%). If one is to compare the photoaffinity labeling of RNA polymerase using different DNA templates (which make different amounts of various RNA chains), quantitative measurements will be necessary.

On the other hand, cutting and counting polyacrylamide gel bands is a rather crude way to extract the information they contain. The 3-mm width of the gel slices is too large to take full advantage of the resolution afforded by electrophoresis but narrow enough to produce hundreds of slices to be counted. The slicing procedure appears to be the limiting factor in our current methodology; automated methods would be expected to produce less scatter in the final results (Anderson et al., 1987).

In our determination of percent yields, the variation from experiment to experiment usually was larger than the error expected from the statistics of radioactive counting. Only when the yields dropped below 0.5% did the statistics of radioactive counting become important. In Figure 4, the results of two independent experiments are shown to allow the reader to judge the precision of the data.

A particular complication observed in these experiments with poly[d(A-T)] may been seen in Figure 3. Near the gel band corresponding to a particular terminated transcript, one often finds a band for the unterminated transcript with the same sequence (e.g., one molecule ends with cordycepin and the other with adenosine). Also, the addition of a U residue changes the electrophoretic mobility less than an A; each A lane contains small amounts of RNAs that end in U and vice versa. It is difficult to accurately exclude the extra bands present in the lanes. Further, light bands that are clearly evident on the film in Figure 3, such as the longer lengths of RNA cleaved from the DNA, contain too few counts for brief measurements in a scintillation counter. A relevant discussion of the problems associated with quantitating closely spaced spots in two-dimensional gels has been presented by Bossinger et al. (1979).

Labeling by Longer Transcripts. For the low but detectable photolabeling of σ in poly[d(A-T)] transcription complexes by transcripts >12 nucleotides long, two possible explanations can be advanced. (1) It may be that σ is still completely associated with the complexes at these longer lengths but the surroundings of the 5′ end of the transcript consist mainly of amino acid side chains within subunits β and β', rather than σ. (2) Another possibility is that due to the peculiar properties of poly[d(A-T)], some of the nascent transcripts form hairpin loops with their 5′ and 3′ ends base-paired to DNA. It is well-known that the enzyme has the ability to transcribe short pieces of poly[d(A-T)] in a continuous fashion that allows the RNA transcripts to become much longer than the DNA template, yet retain the primary contacts that are characteristic of the preinitiation complex (Chamberlin & Berg, 1962; Khorana, 1965). This unusual transcription from poly[d(A-

T)] appears to be the most plausible explanation for the longer RNA lenghts labeling the σ subunit. Such results have not been observed on other templates (Bernhard & Meares, 1986a).

Figure 3 shows that the DNA labeling also continues beyond RNA lengths of 12–13 pairs on poly[d(A-T)]. That has not been seen on the T7 A1 or λ P_R template using the same photoprobe (Hanna & Meares, 1983b; Bernhard & Meares, 1986). The DNA labeling pattern on poly[d(A-T)] approximately parallels that of σ and may also be due to the continuous transcription discussed in the previous paragraph. Another possibility is that the 5′ ends of some fraction of the RNA transcripts do not separate from the poly[d(A-T)] template strand as they pass through the transcription complex.

The results in figures 3 and 4 also make it evident that the 5′-azides on transcripts shorter than seven nucleotides do not label σ. This supports the hypothesis that σ is located near one side of the growing RNA/DNA duplex and that the labeling only occurs on σ as the 5′ end of the transcript passes across this side. It is also clear that the path of nascent RNA through the transcription complex is not the same for different DNA templates. Photoaffinity labeling of the σ subunit appears to be most sensitive to the nature of the DNA.

For *E. coli* RNA polymerase/T7 A1 transcription complexes, the work of Shimamoto et al. (1986) supports the idea of early σ release. This group analyzed the release of σ as a function of time and proposed a two-step model that involves a fast triggering step (<1 s) followed by a slow dissociation (mean lifetime ≈5 s). By analyzing the transcription complexes formed under different GTP concentrations and different incubation times, their results indicate no relationship between the transcript length and the slow dissociation step of σ on the T7 A1 promoter. However, the results of Wu et al. (1975) indicate that the fast triggering step does not take place before the formation of the first two phosphodiester bonds. Shimamoto et al. estimate that the triggering step probably occurs before the transcript is four to six nucleotides long. These results are consistent with the observations of Hanna and Meares (1983a,b) that σ is not photolabeled significantly by transcripts longer than a trinucleotide when the T7 A1 promoter is used.

σ is evidently released quantitatively from transcription complexes on some DNAs, such as T7 A1 (Shimamoto et al., 1986) or lac UV5 (Straney & Crothers, 1987). However, on poly[d(A-T)], it appears that σ release can occur over a range of RNA lengths rather than at a single, unique point during transcription and there may be some transcription complexes that never release σ. This may apply to other DNA templates as well. It is possible that the molecular interactions leading to σ release are analogous to those which occur at promoter or terminator sites, in that they have a range of strengths. It is well-known that different promoter sites on DNA lead to the initiation of transcription with widely different efficiencies (McClure, 1985). Also, studies of the termination of transcription show that it does not occur with 100% efficiency at most termination sites (Chamberlin et al., 1987).

A point of interest from these studies on poly[d(A-T)] becomes evident when one looks at the total yields of photoaffinity labeling in the transcription complex, rather than at the individual subunits. A drastic increase occurs between transcript lengths of 11 nucleotides, at which the overall yield is 22%, and 14 nucleotides, at which the overall yield is 64%. This indicates a major change in the local environment around the 5′ terminus of the transcript in a majority of the complexes. It is dominated by the movement of the 5′-azide into a region of greater contact with β and β', and less contact with σ. Over a similar range of RNA lengths, Levin et al. (1987) have observed (using natural DNA promoters) that the ternary transcription complex becomes stable and ceases abortive initiation of RNA chain synthesis.

The suggestion by Bernhard and Meares (1986a), that the early release of σ from the T7 A1 transcription complex could be mainly controlled by the bridging hydrogen bond donor/acceptor pattern within the major groove of the RNA/DNA duplex, is not supported by the results presented here. Considering only the hydrogen-bond donors and acceptors that bridge the two nucleic acid strands, the poly[d(A-T)] transcript sequence has the same property of alternating hydrogen-bond donors and acceptors as T7 A1. However, the labeling pattern seen here for poly[d(A-T)] more closely resembles that observed for the λ P_R complex than T7 A1.

Since interaction between σ and the nucleic acids evidently is not dominated by contacts with the bridging hydrogen-bond donors and acceptors in the major groove of the RNA/DNA hybrid, the source of the key interaction remains to be determined. Several possibilities exist. (1) A particular set of contacts with the bridging hydrogen-bond donors and acceptors in the major groove of the RNA/DNA hybrid is necessary but not sufficient, so that the actual requirements are more complex; (2) the release of σ may depend on sequences up- or downstream from the transcription start site; (3) the key interaction involves the nontemplate DNA strand, rather than the RNA/DNA hybrid [note that this is known to contact σ (Simpson, 1979; Park et al., 1982a,b)]; (4) other structural parameters such as groove width or minor groove properties like the presence or absence of the 2-amino group of G or a homopurine stretch in the RNA/DNA hybrid are crucial.

Further quantitative data are being gathered on the natural templates T7 A1 and λ P_R, as well as on recombinant constructs, that will allow us to investigate the role of the promoter region vs the early transcript region on the pattern of σ photolabeling (T. Stackhouse, A. Wieland, M. J. Chamberlin, and C. F. Meares, unpublished results). By quantitatively analyzing the photo-cross-linking of the various subunits with the different templates, we hope to obtain a more detailed view of the molecular interactions taking place during transcription.

ACKNOWLEDGMENTS

We thank Susan Bernhard, Jeffrey Marx, and Blaine Bartholomew for helpful comments and discussions.

Registry No. N_3RSpApU, 85977-19-1.

REFERENCES

Anderson, H. L., Puck, T. T., & Shera, E. B. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4749–4753.

Bayley, H., & Knowles, J. R. (1977) *Methods Enzymol. 46*, 69–114.

Bayley, H., & Staros, J. V. (1984) in *Azides and Nitrenes* (Scriven, E. F., Ed.) pp 433–490, Academic, Orlando, FL.

Bernhard, S. L., & Meares, C. F. (1986a) *Biochemistry 25*, 5914–5919.

Bernhard, S. L., & Meares, C. F. (1986b) *Biochemistry 25*, 6397–6404.

Bossinger, J., Miller, M. J., Vo, K.-P., Geiduschek, E. P., & Xuong, N.-H. (1979) *J. Biol. Chem. 254*, 7986–7998.

Burgess, R. R. (1976) in *RNA Polymerase* (Losick, R., & Chamberlin, M., Eds.) pp 69–100, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Burgess, R. R., & Jendrisak, J. J. (1975) *Biochemistry 14*, 4634–4638.

Burgess, R. R., Travers, A. A., Dunn, J. J., & Bautz, E. K. F. (1969) *Nature (London) 221*, 43–46.

Burton, Z., Burgess, R. R., Lin, J., Moore, D., Holder, S., & Gross, C. A. (1981) *Nucleic Acids Res. 9*, 2889–2903.

Carpousis, A. J., & Gralla, J. D. (1980) *Biochemistry 19*, 3245–3253.

Chamberlin, M. J. (1974) *Annu. Rev. Biochem. 43*, 721–775.

Chamberlin, M. J., & Berg, P. (1962) *Proc. Natl. Acad. Sci. U.S.A. 48*, 81–93.

Chamberlin, M. J., Arndt, K. M., Briat, J. F., Reynolds, R. L., & Schmidt, M. C. (1987) in *RNA Polymerase and the Regulation of Transcription* (Reznikoff, W. S., Burgess, R. R., Dahlberg, J. E., Gross, C. A., Record, M. T., Jr., & Wickens, M. P., Eds.) pp 347–356, Elsevier, New York.

Gentry, D. R., & Burgess, R. R. (1986) *Gene 48*, 33–40.

Greenblatt, J., & Li, J. (1981) *Cell (Cambridge, Mass.) 24*, 421–428.

Greenblatt, J., Horwitz, R. J., & Li, J. (1987) in *RNA Polymerase and the Regulation of Transcription* (Reznikoff, W. S., Burgess, R. R., Dahlberg, J. E., Gross, C. A., Record, M. T., Jr., & Wickens, M. P., Eds.) pp 357–366, Elsevier, New York.

Hanna, M. M., & Meares, C. F. (1983a) *Biochemistry 22*, 3546–3551.

Hanna, M. M., & Meares, C. F. (1983b) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4238–4242.

Hansen, U. M., & McClure, W. R. (1980) *J. Biol. Chem. 255*, 9564–9570.

Khorana, H. G. (1965) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 24*, 1473–1487.

Knowles, J. R. (1971) *Acc. Chem. Res. 5*, 155–160.

Levin, J. R., Krummel, B., & Chamberlin, M. J. (1987) *J. Mol. Biol. 196*, 85–100.

Lewin, B. (1983) *Genes*, Chapter 10, Wiley, New York.

Lowe, P. A., Hager, E. A., & Burgess, R. R. (1979) *Biochemistry 18*, 1344–1352.

McClure, W. R. (1985) *Annu. Rev. Biochem. 54*, 171–204.

Ovchinnikov, Yu. A., Lipkin, V. M., Modyanov, N. N., Chertov, O. Yu., & Smirnov, Yu. V. (1977) *FEBS Lett. 76*, 108–111.

Ovchinnikov, Yu. A., Monastyrskaya, G. S., Gubanov, V. V., Guryev, S. O., Chertov, O. Yu., Modyanov, N. N., Grinkevich, V. A., Makarova, I. A., Marchenko, T. V., Polonikova, I. N., Lipkin, V. M., & Sverdlov, E. D. (1981) *Eur. J. Biochem. 116*, 621–629.

Ovchinnikov, Yu. A., Monastyrskaya, G. S., Gubanov, V. V., Guryev, S. O., Salomatina, I. S., Shuvaeva, T. M., Lipkin, V. M., & Sverdlov, E. D. (1982) *Nucleic Acids Res. 10*, 4035–4044.

Park, C. S., Hillel, Z., & Wu, C.-W. (1982a) *J. Biol. Chem. 257*, 6944–6949.

Park, C. S., Wu, F. Y.-H., & Wu, C.-W. (1982b) *J. Biol. Chem. 257*, 6950–6956.

Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804–808.

Shimamoto, N., Kamigochi, T., & Utiyama, H. (1986) *J. Biol. Chem. 261*, 11859–11865.

Siebenlist, U., Simpson, R. B., & Gilbert, W. (1980) *Cell (Cambridge, Mass.) 20*, 269–281.

Simpson, R. B. (1979) *Cell (Cambridge, Mass.) 18*, 277–285.

Straney, D. C., & Crothers, D. M. (1987) *J. Mol. Biol. 193*, 267–278.

Travers, A. A., & Burgess, R. R. (1969) *Nature (London) 222*, 537–540.

von Hippel, P. H. (1979) in *Biological Regulation and Development* (Goldberger, R. F., Ed.) Vol. 1, pp 279–347, Plenum, New York.

von Hippel, P. H., Bear, D. G., Winter, R. B., & Berg, O. G. (1982) in *Promoters: Structure and Function* (Rodriguez, R. L., & Chamberlin, M. J., Eds.) pp 3–33, Praeger, New York.

Wu, C.-W., Yarbrough, L. R., Hillel, Z., & Wu, F. Y.-H. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3019–3023.

Wu, G. J., & Bruening, G. E. (1979) *Virology 46*, 596–612.