A C-Nucleotide Base Pair: Methylpseudouridine-Directed Incorporation of Formycin Triphosphate into RNA Catalyzed by T7 RNA Polymerase†

Joseph A. Piccirilli,‡ Simon E. Moroney,§ and Steven A. Benner*

Laboratory for Organic Chemistry, Swiss Federal Institute of Technology, 8092 Zurich, Switzerland

Received March 14, 1991; Revised Manuscript Received July 11, 1991

ABSTRACT: With templates containing 2′-deoxy-1-methylpseudouridine ($d^{m}\Psi$), T7 RNA polymerase catalyzes the incorporation of either adenosine triphosphate (ATP) or formycin triphosphate (FTP) into a growing chain of RNA with the same efficiency as with templates containing thymidine (dT). In each case, the overall rate of synthesis of full-length products containing formycin is about one-tenth of the rate of synthesis of analogous products containing adenosine. Analysis of the products of abortive initiation shows that incorporation of FMP into the growing oligonucleotide by T7 RNA polymerase is more likely to lead to premature termination of transcription than is incorporation of AMP. Nevertheless, the results demonstrate that T7 RNA polymerase tolerates the formation of a C-nucleotide transcription complex in which the nucleoside bases on both the template and the incoming nucleotide are joined to the ribose by a carbon–carbon bond. This result increases the prospects for further expanding the genetic alphabet via incorporation of new base pairs with novel hydrogen-bonding schemes (Piccirilli et al., 1990).

Life on earth seems to have passed through several episodes, including one in which ribonucleic acid (RNA) was the sole genetically encoded component of biological catalysis (Rich, 1962; White, 1976; Visser & Kellogg, 1978; Orgel, 1986; Benner et al., 1989). Examples of RNA molecules having catalytic activity support this view (Cech & Bass, 1986), and models that presume the existence of an "RNA world" rationalize many of the metabolic details of contemporary organisms (Benner et al., 1989). The most plausible of these models suggests that RNA catalyzed many types of reactions in the "RNA world" prior to the emergence of proteins synthesized by translation of a genetically encoded message (Benner et al., 1987, 1989).

Nevertheless, on chemical grounds, the catalytic potential of RNA is severely limited by the small number of building blocks and the correspondingly limited functionality available to RNA, especially when compared with proteins. This limitation has slowed progress in the laboratory toward developing self-replicating molecules from natural RNA (e.g., the type 1 self-splicing intron) (Doudna et al., 1991; Robertson & Joyce, 1990) or other systems (von Kiedrowski, 1986).

The limitation on catalytic power imposed by the small number (four) of nucleoside building blocks in natural RNA can be overcome in principle by incorporating additional replicatable building blocks into nucleic acids (Benner et al., 1987). These additional "letters" in the genetic alphabet are possible because the Watson–Crick base-pairing geometry allows six different hydrogen-bonding patterns, corresponding to 12 independently replicatable nucleoside bases (see Figure 1; Benner et al., 1987; Switzer et al., 1989; Piccirilli et al., 1990). An experimental goal of this laboratory has been to use these additional building blocks to create self-replicating RNA molecules.

In practice, however, DNA and RNA polymerases must accept a new base pair before oligonucleotides containing it are easily accessible. Although DNA and RNA polymerases have now been found that accept base pairs between isocytidine and isoguanosine and between xanthosine (X) and diaminopyrimidine (κ)[1] (Switzer et al., 1989; Piccirilli et al., 1990), polymerases may not accept structures that differ still more from those of the natural nucleotides.

This is especially true with respect to one structural feature important to many of the new base pairs: a carbon–carbon bond joining the ribose and the nucleoside base. A C-nucleoside is often desired both for the pyrimidine analogue and for the purine analogue. Three of the pyrimidine analogues in Figure 1 require glycosidic C–C bonds to achieve the necessary disposition of hydrogen-bonding groups. The use of C-nucleoside purine analogues may be advantageous to ensure a purine–sugar linkage stable to acid-catalyzed glycosidic cleavage. However, replacing the normal carbon–nitrogen glycosidic bond by a C–C bond may alter the conformation of the ribose ring due to different steric and stereoelectronic interactions between segments of the molecule. Some of the conformational features of C-glycosides in the crystal may be due to such effects (Giranda et al., 1988; Prusiner et al., 1973; Koyama et al., 1974, 1976). As the conformation of the ribose ring and the conformation of the nucleoside are both possible recognition elements for polymerases, natural polymerases may not accept new base pairs made up of two C-nucleosides.

Neither of the base pairs that have been examined so far (Switzer et al., 1989; Piccirilli et al., 1990) have incorporated

† Supported by the Swiss National Science Foundation, the ETH, and Sandoz AG.

* Author to whom correspondence should be addressed.

‡ Present address: Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309.

§ Present address: Department of Pharmacology, Cambridge University, Cambridge, England CB21QJ.

[1] Abbreviations: $d^{m}\Psi$, 2′-deoxy-1-methylpseudouridine; F, formycin A; FTP, formycin triphosphate; FMP, formycin monophosphate; κ, 3-β-D-ribofuranosyl-2,6-diaminopyrimidine; X, xanthosine; Ψ, pseudouridine; DMT, 4,4′-dimethoxytrityl; HPLC, high-performance liquid chromatography; TEAAc, triethylammonium acetate; tlc, thin-layer chromatography; EtOAc, ethyl acetate; T7T1, T7 RNA polymerase promoter–template complex containing thymidine; T7Ψ1; T7 RNA polymerase promoter–template complex containing pseudouridine; DTT, dithiothreitol; BSA, bovine serum albumin; EDTA, ethylenediaminetetraacetic acid; NMR, nuclear magnetic resonance; NTP, nucleoside triphosphate.

0006-2960/91/0430-10350$02.50/0

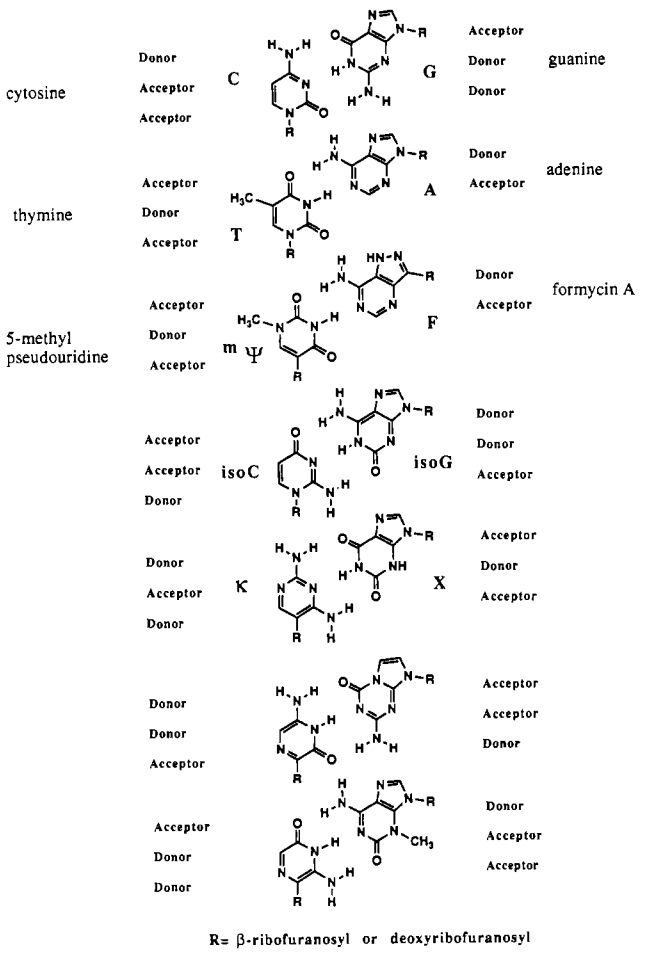

FIGURE 1: Six mutually exclusive hydrogen-bonding patterns that are consistent with the standard geometry of the Watson–Crick base pair. C-glycosides are found in the last three base pairs. The hydrogen-bonding pattern of the pseudouracil–formycin base pair is identical with that of the T–A base pair.

two C-nucleotides opposite one another in a DNA duplex (Benner et al., 1987). Before attempting to find conditions under which an RNA polymerase would accept a base pair having *both* a new hydrogen bonding scheme *and* two C-nucleotides as components, we wished to learn whether polymerases can accept a new base pair between two C-nucleotides that have a "natural" hydrogen-bonding scheme. In addition to separating experimentally the two novel structural features of many of the new base pairs, such studies should provide insight into the way in which polymerases recognize their substrates.

Pseudouridine (Ψ) (Hori et al., 1964) and formycin A (F) (Cohn, 1960) (Figure 1) are two naturally occurring C-nucleosides that can form a base pair having the same hydrogen-bonding pattern as the uridine–adenosine pair (Ward & Reich, 1968). Although the pioneering studies of Ward et al. (1969) showed that the RNA polymerase from *Escherichia coli* would accept formycin triphosphate as a substrate, no work has been done with any of the modern, high yielding RNA polymerases used for the in vitro synthesis of RNA (e.g., T7 RNA polymerase) (Milligan et al., 1987). As a means of preparing milligram amounts of oligonucleotides containing modified bases for subsequent studies, we have investigated whether C-nucleotides are substrates for T7 RNA polymerase. Furthermore, T7 RNA polymerase produces "abortive" transcription products that can provide information regarding the phase(s) of transcription (incorporation, translocation, elongation, processive continuation) perturbed by the presence of the analogue (Martin et al., 1988; Moroney & Piccirilli, 1991).

We report here that 2′-deoxy-1-methylpseudouridine ($d^{m}\Psi$) in a DNA template directs, with essentially the same efficiency as thymidine (dT), the incorporation of ATP into an RNA transcript, catalyzed by T7 RNA polymerase. Formycin triphosphate (FTP) can substitute for ATP, but at a significantly reduced rate. Analysis of products formed by abortive initiation revealed that FMP incorporation is more likely to lead to premature termination of transcription than is incorporation of AMP. Nevertheless, the results demonstrate that T7 RNA polymerase catalyzes the formation of a C-nucleotide transcription complex in which a template containing the C-nucleoside 2′-deoxy-1-methylpseudouridine residue directs the incorporation of the C-nucleoside formycin at a specific site.

Materials and Methods

Chemicals and Reagents. Pyridine was distilled from calcium hydride and stored over 3-Å molecular sieves. Dichloromethane was distilled first from phosphorus pentoxide and then calcium hydride. Solvents were purchased in bulk and distilled over calcium chloride. Pseudouridine (β-anomer) was from Sigma. 4,4′-Dimethoxytrityl chloride was from Fluka. Bis(diisopropylamino)methoxyphosphine was prepared according to published procedures (Caruthers et al., 1984). Nucleoside triphosphates were purchased from Pharmacia, dissolved in water to a concentration of 40 mM, adjusted to pH 8.1 with 0.5 M Tris, and stored at −20 °C. [α-^{32}P]CTP (3000 Ci/mmol) and [γ-^{32}P]GTP (10 Ci/mmol) were from Amersham International.

Formycin A triphosphate (FTP), containing at most 2% contaminating ATP, was purchased from Calbiochem and purified by HPLC on a Mino RPC C2/C18 column (Pharmacia) using a 20-min linear gradient from 0.1 M triethylammonium acetate, pH 7.0 (TEAAc) to 10% acetonitrile in TEAAc. FTP was dissolved in water to a concentration of 40 mM (based on an extinction coefficient of 10300 M^{-1} cm^{-1} (Maeliche et al., 1974) and adjusted to pH 8.1, as described above.

Enzymes. Snake venom phosphodiesterase (*Crotalus durissus*, 1.5 units/mg), and RNAsin ribonuclease inhibitor were purchased from Boehringer Mannheim. Bacterial alkaline phosphatase (*E. coli*, 500 units/mL) was from Calbiochem. T4 polynucleotide kinase was from New England BioLabs, Inc. T7 RNA polymerase (175000 units/mg, 70000 units/mL) was from Pharmacia.

Synthesis

2′-Deoxy-1-methylpseudouridine (**1**). This compound was synthesized from 1-methylpseudouridine according to the method of Chu et al. (1977): mp 158–160 °C (lit. mp 158–160 °C); UV (H_2O) λ_{max} 271 nm (ϵ 9000), λ_{min} 236 nm (ϵ 1600), 260 nm (ϵ 7800), 280 nm (ϵ 6500). The ^{1}H NMR spectrum was identical with that reported in the literature.

2′-Deoxy-5′-O-(4,4′-dimethoxytrityl)-1-methylpseudouridine ($d^{m}\Psi$) (**2**). 2′-Deoxy-1-methylpseudouridine **1** (213 mg, 0.88 mmol) was dissolved in dry pyridine (4.4 mL) and treated with triethylamine (1.4 equiv, 1.7 mL) and (dimethylamino)pyridine (0.5 equiv, 5.4 mg). 4,4′-Dimethoxytrityl chloride (1.2 equiv, 369 mg) was added, and the resulting dark brown solution was stirred under argon at room temperature. The reaction was monitored by tlc (silica, 10% $MeOH/CH_2Cl_2$, R_f = 0.4). After 2 h, more DMTrCl (0.24 equiv, 88 mg) was added and the mixture was treated with additional triethylamine (34 μL). After 2 h more, tlc showed

that all of the starting material had been consumed. The mixture was then diluted with diethyl ether (50 mL) and extracted with water (5 mL). The layers were separated, and the aqueous layer was washed with ether (50 mL). The combined organic layers were dried (Na_2SO_4), and solvent was removed in vacuo. The residue was purified by flash chromatography (silica gel) with use of a stepwise gradient from 100% CH_2Cl_2 to 3% MeOH in CH_2Cl_2. Combination of fractions containing product followed by removal of solvent in vacuo gave **2** as a white foam (359 mg, 75%). The compound was used directly in subsequent reactions. An analytical sample was obtained as white plates by recrystallization (benzene/hexane): mp 75 °C; ^{1}H NMR ($CDCl_3$) δ 9.1 (br, 1 H, NH), 7.42 (m, 3 H, Ph), 7.30 (m, 4 H, Ph), 6.83 (d, 4 H, Ph), 5.11, (t, 1 H, H-1′), 4.44 (m, 1 H, H-3), 4.04 (m, 1 H, H-4), 3.80 (s, 6 H, OCH_3), 3.29 (d, 2 H, H-5′, H-5″), 3.14 (s, 3 H, CH_3), 2.48 (m, 1 H, H-2′), 1.95 (1 H, H-2″); IR (KBr) 3420 (br), 2960, 2930, 1670, 1600, 1510, 1250, 1180, 1030, 830, 580 cm^{-1}; MS, *m/e* (rel intensities) 544 (<1%, M^+), 303 (39.7, trityl), 153 (27.6, B-CH_2CH_2). Anal. Calcd for $C_{31}H_{32}N_2O_7$: C, 68.38; H, 5.88; N, 5.15. Found: C, 68.12; H, 5.79; N, 4.98.

3′-O-[[(Diisopropylamino)methoxy]phosphino]-5′-O-(4,4′-dimethoxytrityl)-2′-deoxy-1-methylpseudouridine (**3**). Protected nucleoside **2** (216 mg, 0.4 mmol) was twice dissolved in 10% CH_2Cl_2/pyridine and evaporated to dryness in vacuo to remove residual water. Toluene was added, and the mixture was again evaporated to dryness in vacuo. The flask containing **2** was equipped with a septum, vented with a needle, and allowed to stand in a desiccator overnight under vacuum at −20 °C. Diisopropylammonium tetrazolide (0.5 equiv, 34.3 mg) was added, and the flask was evacuated and refilled with Ar. The mixture was dissolved in CH_2Cl_2 (2 mL) and treated with bis(diisopropylamino)methoxyphosphine (1.1 equiv, 140 μL) at room temperature. Monitoring by tlc showed the formation of product (R_f = 0.5; 10% triethylamine, 45% CH_2Cl_2, 45% EtOAc) with consumption of starting material. After 1 h, the reaction mixture was poured into a saturated aqueous solution of sodium bicarbonate, and the organic layer was washed with brine and dried over sodium sulfate. Evaporation of the solvent gave a white foam, which was taken up in toluene (1 mL) and precipitated by addition to hexane (−78 °C) with rapid stirring. The resulting suspension was allowed to stir for 10 min at −78 °C, filtered cold, and dried under vacuum to afford 226 mg (80%) of **3**, which was stored under argon at −20 °C. MS, *m/e* (rel intensities) 705 (1, M^+) 527 (52.6, [M^+–OPOMe(iPr)$_2$]); ^{31}P NMR δ 148.5, 149.03.

Synthesis of Oligonucleotides. All oligodeoxynucleotides were synthesized according to standard cyanoethyl phosphoramidite chemistry on an Applied Biosystems Model 380B DNA synthesizer. The dimethoxytrityl group of the 5′-nucleoside was left on the oligodeoxynucleotide to allow purification via trityl selection reversed-phase HPLC. A 20-fold excess of d$^m\Psi$ phosphoramidite was used instead of the usual 10-fold excess, and the phosphoramidite and coupling reagents were delivered twice. The coupling yield at the step when d$^m\Psi$ was added was 80–90%.

Purification of Oligonucleotides. All oligonucleotides were purified by reversed-phase HPLC on a Vydac C4 semipreparative column using a 20-min linear gradient from 15 to 30% acetonitrile in 0.1 M TEAAc. The purified oligonucleotides were detritylated by treatment with 80% HOAc. The purity of the products was confirmed by 5′-end-labeling with [γ-^{32}P]ATP and T4 polynucleotide kinase and analysis of the product by denaturing gel electrophoresis (12% polyacrylamide, 7 M urea). Template–promoters were annealed by heating a 1:1 mixture (based on an average extinction coefficient of 10^4 M^{-1} cm^{-1}) of the T7 RNA polymerase promoter strand (T7P) and either T7T1 or T7Ψ1, at 70 °C for 3 min, and then allowing the solution to cool slowly to room temperature.

Enzymatic Digestion. Oligonucleotides (1.0 OD units) were incubated in 50 mM Tris-HCl (pH 8.0)/10 mM $MgCl_2$ with snake venom phosphodiesterase (0.2 unit) and bacterial alkaline phosphatase (0.5 unit) in a total volume of 40 μL at 37 °C for 5 h. The reaction mixture was passed through a Centricon-10 membrane to remove protein, and the deoxynucleosides were analyzed by HPLC on a Mino RPC C2/C18 reversed-phase column (Pharmacia) using a 20-min linear gradient from TEAAc to 25% acetonitrile in TEAAc. The retention times observed under these conditions were (dC) 5.86 min, (d$^m\Psi$) 7.07 min, (dG) 7.83 min, (dT) 8.07 min, and (dA) 10.11 min. The α-anomer of 2′-deoxy-1-methylpseudouridine was not detected.

Polyacrylamide gel (20%) electrophoresis was on slab gels (48 × 16 × 0.06 cm) containing 7 M urea (Maniatis & Efstratiadis, 1980). Gels were prerun (500 V for 2–3 h) and run at 400–600 V for 6–8 h after application of samples. Bands were visualized by autoradiography using Kodax X-Omat film. For quantification, bands were excised, and the RNA was liberated for scintillation counting by treatment with Protosol (Du Pont; 0.3 mL, 35 °C, 4-6 h). Scintillation fluid (Kontron Kontrogel, 10 mL) and acetic acid (2 drops) were then added to the Protosol solution, and radioactivity was determined with use of a Kontron Betamatic II liquid scintillation counter. Under these conditions, extraction of RNA from the gel was found to be quantitative, and scintillation counting efficiency was between 75 and 80%.

Transcription Reactions. Incubation mixtures contained $MgCl_2$ (20 mM), spermidine (1 mM), DTT (5 mM), RNasin (50 units), nucleoside triphosphates (2.5 mM each), Triton X-100 (0.01%), template TpT1 or TpΨ1 (4.0 μg/mL), BSA (50 μg/mL), either [γ-^{32}P]GTP or [α-^{32}P]CTP (5–10 μCi), and T7 RNA polymerase (140 units) at 38 °C in 20 μL of 40 mM Tris-HCl buffer (pH 8.1). For kinetic work, the reaction was initiated by the addition of a mixture of nucleoside triphosphates, preequilibrated at 38 °C, to the other components at the same temperature, thereby permitting preassociation of enzyme and template. Aliquots (4 μL) were removed from the reaction at time intervals and quenched by addition of an equal volume of sample buffer (7 M urea, 89 mM Tris–borate, 2 mM EDTA, 0.02% xylene cyanol, 0.02% bromophenol blue) and heating at 70 °C for 5 min. The quenched aliquots were then stored on ice prior to analysis by electrophoresis.

Rates and yields of full-length products were determined from reactions in which [γ-^{32}P]GTP provided the label. In this way, all reaction products were visualized by autoradiography and quantitated by excision from the gel and scintillation counting.

Results

Incorporation of 2′-Deoxy-1-methylpseudouridine (d$^m\Psi$) into DNA. Pseudouridine epimerizes under both acidic and basic conditions to yield a mixture of α- and β-anomers (Chambers et al., 1963). As the mechanism for epimerization involves deprotonation at N^1, 2′-deoxy-1-methylpseudouridine (d$^m\Psi$), a derivative lacking a proton at this center, and an isostere of thymidine, was used in these experiments. Chromatographic and ^{1}H NMR data (not shown) confirmed that d$^m\Psi$ was unaffected by treatment either with trichloroacetic

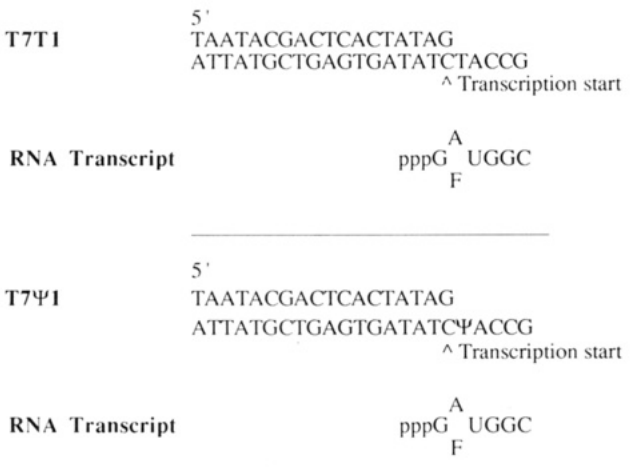

FIGURE 2: Structures of the templates T7T1 and T7Ψ1 and the RNA transcripts that they encode.

acid in dichloromethane (30 min) or with 29% NH_4OH (17 h, 60 °C) overnight, conditions used for the synthesis of DNA. The nucleoside was then converted to a form suitable for automated DNA synthesis by use of standard chemistry. Thus, treatment of $d^m\Psi$ with 4,4′-dimethoxytrityl chloride in pyridine yielded 5′-(dimethoxytrityl)-N^1-methylpseudouridine in 75% yield. Reaction with bis(diisopropylamino)methoxyphosphine in the presence of a weak acid catalyst (Caruthers et al., 1984) yielded the corresponding phosphoramidite, whose authenticity and purity were confirmed by ^{31}P NMR.

Templates of the "overlapping" type (Milligan et al., 1987) incorporated a double-stranded T7 RNA polymerase promoter followed by a single-stranded coding sequence designed to produce hexanucleotide transcripts. The sequences of the templates T7T1 and T7Ψ1 are given in Figure 2, the only difference between the two being at the second transcribed position. Yields for incorporation of $d^m\Psi$ on a 1-μmol scale were approximately 90%, as determined by assay of the subsequently released dimethoxytrityl cation.

After being purified by HPLC, the synthetic oligomers were homogeneous by 5′-end-labeling and analysis by polyacrylamide gel electrophoresis. The presence of 2′-deoxy-1-methylpseudouridine in the oligonucleotide T7Ψ1 was verified by digestion of a specimen with snake venom phosphodiesterase and alkaline phosphatase, followed by analysis of the resulting mixture by reversed-phase HPLC. Nucleosides were present in the expected ratio and none of the α-anomer of 2′-deoxy-1-methylpseudouridine was present (data not shown).

The $d^m\Psi$-Containing Template Is Accepted by T7 RNA Polymerase. Transcription products obtained from template T7T1 upon protracted incubation yielded a hexanucleotide as the major product, together with both shorter and longer labeled products (Figure 3, lane 2). The additional bands observed at a longer time are consistent with the known ability of T7 RNA polymerase to add nonencoded residues at the 3′-end of transcripts (Milligan et al., 1987) and, evidently, some misincorporation of [α-^{32}P]CTP into the shorter transcripts. These additional products were absent in shorter incubations.

The relative abilities of 2′-deoxy-1-methylpseudouridine and formycin triphosphate to act, respectively, as template and "incoming" nucleotides were compared to those of the corresponding natural nucleotides in parallel experiments. Figure 3 shows results where label was incorporated in the last base of the hexanucleotide, allowing visualization of the full-length products only. Under identical conditions, transcription from templates T7T1 and T7Ψ1 yielded essentially the same

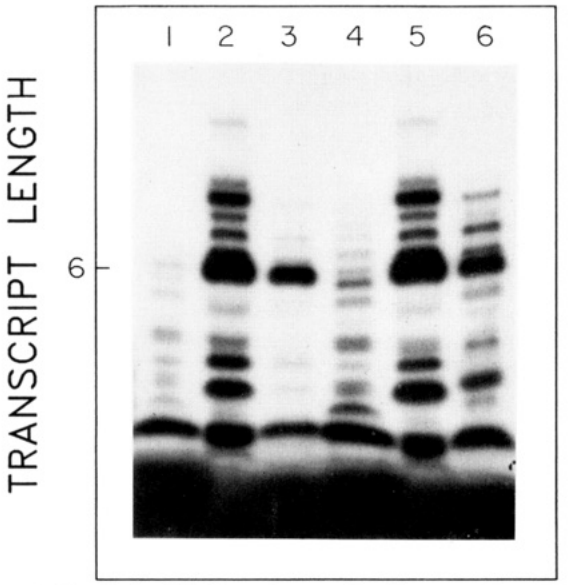

FIGURE 3: Synthesis of full-length transcripts from templates T7T1 and T7Ψ1 by T7 RNA polymerase. Autoradiograph of gel showing products obtained from templates T7T1 (lanes 1–3) and T7Ψ1 (lanes 4–6) in the presence of [α-^{32}P]CTP, GTP, and UTP (lanes 1 and 4) or [α-^{32}P]CTP, GTP, UTP, and ATP (lanes 2 and 5) or [α-^{32}P]CTP, GTP, UTP, and FTP (lanes 3 and 6) in a 4-h transcription reaction.

amounts of full-length products. Further, the initial rates of synthesis of full-length product are roughly the same for templates containing either T or $d^m\Psi$. The close similarity in the results of experiments comparing the two types of templates was observed regardless of which purine analogue (ATP or FTP) was used. Further, in the absence of ATP and FTP, both templates yielded only minor amounts of hexanucleotide (Figure 3, lanes 1 and 4). This observation suggests that the fidelity of transcription was high, with only AMP or FMP being incorporated opposite T and $d^m\Psi$.

However, the substitution of FTP for ATP had a significant effect on the rate of synthesis of full-length product. At short incubation times, the rate of synthesis of full-length product is reduced by 87% (for template T7T1) and by 91% (for template T7Ψ1) when FTP replaces ATP (Figure 5b).

The use of [γ-^{32}P]GTP as a substrate allows incorporation of radioactive label at the beginning of the transcript and visualization of all products formed, including those resulting from premature termination of the transcription reaction (Martin et al., 1988). The amounts of each RNA product formed were measured at 2, 4, 6, and 8 min in four reactions run in parallel, each containing template T7T1 or T7Ψ1 and either ATP or FTP in addition to the other three nucleoside triphosphates. Figure 4 shows an autoradiograph of a typical gel in which the products of the four reactions were analyzed. Prematurely terminated transcripts of lengths 2, 3, 4, and 5 can be seen in addition to the full-length product. The bands were assigned by a series of "minus" transcription experiments and nearest-neighbor analysis (Moroney & Piccirilli, 1991). In each case, products were separated by polyacrylamide gel electrophoresis, excised from the gel, and quantified by scintillation counting. The data for hexanucleotide and total RNA formation as derived from Figure 4 are summarized in Table I and Figure 5.

Figures 4 and 5 again show that the dT- and $d^m\Psi$-containing templates behave similarly. The rate of synthesis of all RNA products (full-length plus abortive) taken together that incorporate FTP is, however, only 40% lower than the rate of incorporation of ATP (Figure 5a). The incorporation of FMP into a growing oligonucleotide chain results in the enzyme

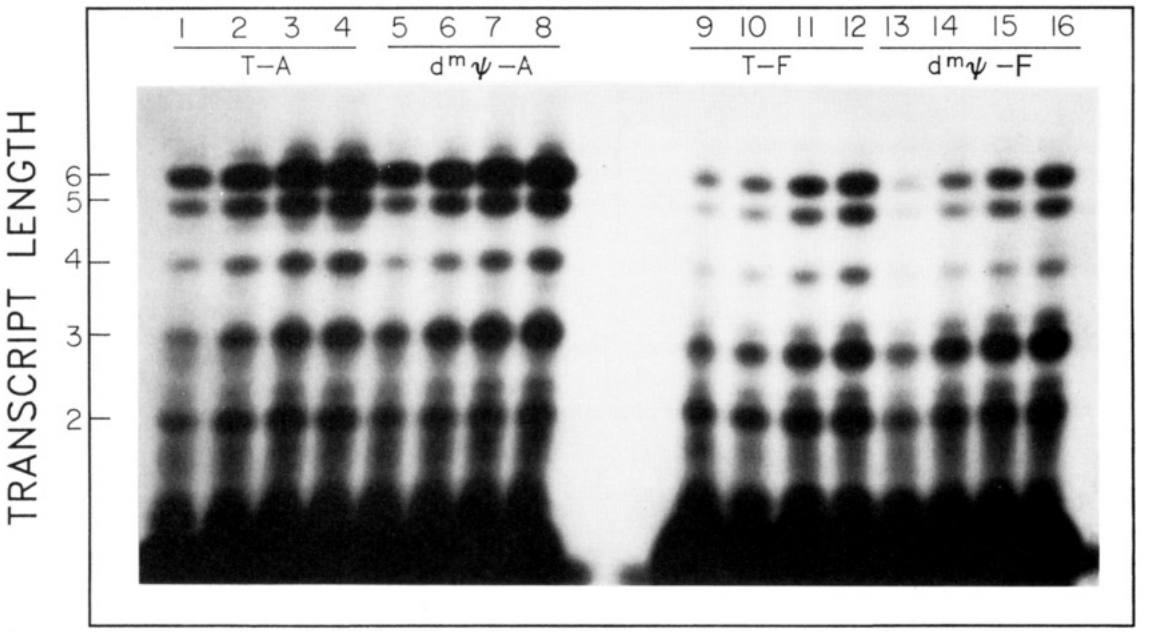

FIGURE 4: Time course of incorporation of [γ-^{32}P]GTP into aborted and full-length products. Autoradiograph of a typical gel in which the products of the four reactions were analyzed. Each reaction contained UTP, CTP, and [γ-^{32}P]GTP. Lanes marked A and F contained ATP and FTP, respectively. Templates are also indicated. Four aliquots corresponding to an incubation time of 2 min (lanes 1, 5, 9, and 13), 4 min (lanes 2, 6, 10, and 14), 6 min (lanes 3, 7, 11, and 15), and 8 min (lanes 4, 8, 12, and 16) for each reaction were electrophoresed. Full-length and prematurely terminated transcripts of length 2, 3, 4, and 5 are seen and assigned by a series of "minus" transcription experiments and nearest-neighbor analysis (Moroney & Piccirilli, 1991). Relative amounts of products, determined by scintillation counting, are summarized in Table I and product time profiles are given in Figure 5.

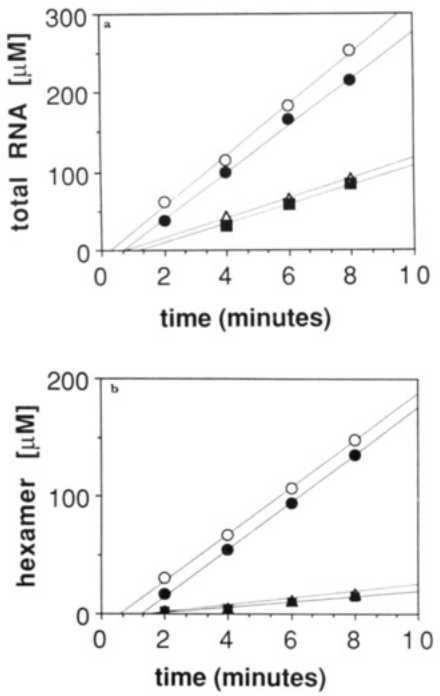

FIGURE 5: Product time curves for RNA synthesis involving elongation with ATP or FTP using templates T7T1 and T7Ψ1. The data are obtained from quantification of the gel in Figure 4. All reactions contained UTP, CTP, [γ-^{32}P]GTP, and the following template–triphosphate combinations: T7T1–ATP (○); T7Ψ1–ATP (●); T7T1–FTP (△); T7Ψ1–FTP (■). Data for the production of total RNA transcripts, including abortive products (panel a), and for full-length product only (panel b) were fitted by a least-squares analysis. The lines do not pass through the origin because of an initial lag phase that has been ascribed to rate-limiting preassociation of enzyme and template (Chamberlin & Ring, 1973).

Table I: Initial Velocities of Formation of Full-Length Product (hexamer) in Transcription Reactions (Data from Figure 5b)

template	N[a]	initial velocity (μM min^{-1})	relative velocity
T7T1	A	19.6	10.3
	F	2.6	1.4
T7Ψ1	A	20.3	10.7
	F	1.9	1.0

[a] Each reaction contained UTP, CTP, [γ-^{32}P]GTP, and NTP, with N as given in the second column. Kinetic data were obtained by quantification of hexanucleotide transcripts formed after 2, 4, 6, and 8 min of reaction time. See Materials and Methods for further details.

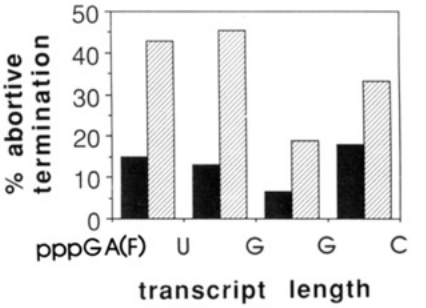

FIGURE 6: Representation that incorporation of formycin A into a growing RNA transcript results in a greater tendency for T7 RNA polymerase to abort synthesis: comparison of abortive tendencies of T7T1-directed transcription in the presence of ATP (solid bars) or FTP (hatched bars). The percentage of transcripts that abortively terminated before the next elongation event is denoted by a bar placed between the two bases. Dividing the amount of abortive product of a given length by the total amount of products of that length or longer gives the percentage of abortive termination at a given position in the transcript (Martin & Coleman, 1988). The values indicated are derived from the data in Table II.

being 3 times more likely (relative to incorporation of AMP) to terminate RNA synthesis before reaching full-length product. This difference in frequency of premature termination is maintained one, two, or three nucleotides downstream from the site of incorporation of AMP or FMP (Figure 6).

DISCUSSION

These results show that T7 RNA polymerase can recognize and accept as a substrate a base pair consisting of two C-

Table II: Amounts of Abortive and Full-Length Transcription Products Observed in Figure 4[a]

NTP	transcript length	sequence	yield (nmol) $T7T_1$	yield (nmol) $T7\Psi_1$
	2	pppGA	0.51	0.40
	3	pppGAU	0.36	0.62
ATP	4	pppGAUG	0.16	0.15
	5	pppGAUGG	0.42	0.40
	6	pppGAUGGC	1.92	2.12
total transcription products			3.37	3.69
	2	pppGF	0.51	0.53
	3	pppGFU	0.31	0.50
FTP	4	pppGFUG	0.07	0.05
	5	pppGFUGG	0.10	0.08
	6	pppGFUGGC	0.21	0.16
total transcription products			1.20	1.31

[a] Reaction conditions were as described under Materials and Methods. The products are the result of a 6-min incubation period at 37 °C.

nucleosides. Because these C-nucleosides form a Watson–Crick base pair with a hydrogen-bonding scheme identical with that found in the natural adenine–thymine pair, these experiments probe the effect of a particular structural perturbation, a pair of carbon–carbon glycosidic linkages, on transcription. Thus, these results provide a preview of work attempting the enzymatic incorporation of base pairs composed of C-nucleosides having novel hydrogen-bonding patterns (Figure 1).

Because T7 RNA polymerase shows no tendency to terminate synthesis at the site containing $d^m\Psi$, it appears that the enzyme cannot distinguish $d^m\Psi$ and dT in the template at any step of transcription. In contrast, the initial rate of synthesis of full-length transcript is reduced 7–11-fold when FTP replaces ATP with either the template containing dT or the template containing $d^m\Psi$. This reduced rate is due to the perturbation of several microscopic kinetic steps. Although the results do not reveal whether FTP is incorporated more slowly than ATP, they do show that once FMP is incorporated, the polymerase is three times more likely to terminate RNA synthesis before reaching full-length product than with ATP, with the incidence of abortion increased approximately equally one, two, and three nucleotides downstream from the site of incorporation of F.

If the impact of the base analogue were simply to misorient the 3′-OH of the unnatural base that is acting as the nucleophile for the elongation reaction, one might expect that the abortive release of the growing RNA chain would occur more frequently at the elongation step following incorporation of FMP than in subsequent steps. In fact, the presence of formycin in the product strand increases the probability that chain synthesis terminates prematurely throughout the transcription experiment reported here. Although the presence of formycin in the growing RNA chain may influence the incorporation of subsequent bases in longer transcripts, its effect is likely to disappear as soon as the polymerase has moved sufficiently far down the template that it is no longer in contact with the modified base. It is possible that the structural perturbation weakens the forces holding the transcription product together with the template in the transcription complex, allowing dissociation of the RNA transcript to compete more effectively with incorporation of successive bases along the chain. It will be interesting to learn whether the extent of abortive cycling can be rigorously correlated with the stability of the duplex synthesized.

Some 20 years ago, Ward and Reich (1968) observed that formycin-containing polymers have abnormal properties that were attributed to the tendency of formycin to adopt the syn conformation. A preference for the syn conformer implies, of course, a less favorable standard double-helical structure, which requires bases in an anti conformation, and therefore provides one possible structural explanation for the results observed here. Alternatively, it is possible that T7 RNA polymerase directly interacts with a functional group present on adenine incorporated into RNA that is distorted in, or absent from, formycin A. The 2′-hydroxyl group (the position of which is distorted by the change in the conformation of the sugar ring in formycin) and the lone pair of electrons on N^7 of both adenine and guanine (replaced by an NH group in formycin A) are two candidates for such a group.

References

Benner, S. A., Allemann, R. K., Ellington, A. D., Ge, L., Glasfeld, A., Leanz, G. F., Krauch, T., MacPherson, L. M., Moroney, S. E., Piccirilli, J. A., & Weinhold, E. G. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 53–63.

Benner, S. A., Ellington, A. D., & Tauer, A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7054–7058.

Caruthers, M. H., Tang, J. Y., & Barone, A. D. (1984) *Nucleic Acids Res. 12*, 4051–4061.

Cech, T. R., & Bass, B. L. (1986) *Annu. Rev. Biochem. 55*, 599–629.

Chamberlin, M. J., & Ring, J. (1973) *J. Biol. Chem. 248*, 2235–2244.

Chambers, R. W., Kurkov, V., & Shapiro, R. (1963) *Biochemistry 2*, 1192–1203.

Chu, C. K., Reichman, U., Watanabe, K. A., & Fox, J. J. (1977) *J. Org. Chem. 42*, 711–714.

Cohn, W. E. (1960) *J. Biol. Chem. 235*, 1488–1498.

Doudna, J. A., Couture, S., & Szostak, J. W. (1991) *Science 251*, 1605–1608.

Giranda, V. L., Berman, H. M., & Schramm, V. L. (1988) *Biochemistry 27*, 5813–5818.

Hori, M., Ito, E., Takita, T., Koyama, G., Takeuchi, T., Umezawa, H. (1964) *J. Antibiot. (Tokyo), Ser. A 17*, 96–99.

Koyama, G., Umezawa, H., & Iitaka, Y. (1974) *Acta Crystallogr. B30*, 1511–1516.

Koyama, G., Nakamura, H., Umezawa, H., & Iitaka, Y. (1976) *Acta Crystallogr. B32*, 813–820.

Maeliche, A., Sprinzl, M., von der Haar, F., Khwaja, T. A., & Cramer, F. (1974) *Eur. J. Biochem. 43*, 617–625.

Maniatis, T., & Efstratiadis, A. (1980) *Methods Enzymol. 65*, 299–305.

Martin, C. T., Muller, D. K., & Coleman, J. E. (1988) *Biochemistry 27*, 3966–3974.

Milligan, J. F., Groebe, D. C., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783–8798.

Moroney, S. E., & Piccirilli, J. A. (1991) *Biochemistry* (preceding paper in this issue).

Orgel, L. (1986) *J. Theor. Biol. 123*, 127–149.

Piccirilli, J. A., Krauch, T., Moroney, S. E., & Benner, S. A. (1990) *Nature 343*, 33–37.

Prusiner, P., Brennan, T., & Sundaralingam, M. (1973) *Biochemistry 12*, 1196–1202.

Rich, A. (1962) in *Horizons in Biochemistry* (Kasha, M., & Pullmann, B., Eds.) pp 103–126, Academic Press, New York.

Robertson, D. L., & Joyce, G. F. (1990) *Nature 344*, 467–468.

Switzer, C., Moroney, S. E., & Benner, S. A. (1989) *J. Am. Chem. Soc. 111*, 8322–8323.
Visser, C. M., & Kellogg, R. M. (1978) *J. Mol. Evol. 11*, 171–187.
von Kiedrowski, G. (1986) *Angew. Chem. 98*, 932–934.
Ward, D. C., & Reich, E. (1968) *Proc. Natl. Acad. Sci. U.S.A. 61*, 1494–1501.
Ward, D. C., Cerami, A., Reich, E., Acs, G., & Altwerger, L. (1969) *J. Biol. Chem. 244*, 3243–3250.
White, H., III (1976) *J. Mol. Evol. 7*, 101–104.

Functional Expression and Characterization of the Interferon-Induced Double-Stranded RNA Activated P68 Protein Kinase from *Escherichia coli*†

Glen N. Barber,‡ Judy Tomita,‡ Ara G. Hovanessian,§ Eliane Meurs,§ and Michael G. Katze*,‡

Department of Microbiology, School of Medicine, University of Washington, Seattle, Washington 98195, and Unité de Virologie and Immunologie Cellulaire, Institut Pasteur, Paris, France

Received June 4, 1991; Revised Manuscript Received July 30, 1991

ABSTRACT: The P68 protein (referred to as P68 on the basis of its molecular weight of 68000 in human cells) is a serine/threonine kinase induced by interferon treatment and activated by double-stranded (ds) RNAs. Although extensively studied, little is currently known about the regulation of kinase function at the molecular level. What is known is that activation of this enzyme triggers a series of events which lead to an inhibition of protein synthesis initiation and may, in turn, play an integral role in the antiviral response to interferon. To begin to understand P68 and its biological functions in the eukaryotic cell, we have expressed the protein kinase in *Escherichia coli* under control of the bacteriophage T7 promoter. In rifampicin-treated cells, metabolically labeled with [^{35}S]methionine and induced by IPTG, the P68 kinase was the predominant labeled product. Further, P68 was recovered from extracts as a fully functional enzyme, shown by its ability to become activated and phosphorylate its natural substrate, the α subunit of eukaryotic protein synthesis initiation factor 2 (eIF-2). Moreover, P68 was phosphorylated in vivo in *E. coli*, providing conclusive evidence that the kinase has the capacity to phosphorylate and activate itself in the absence of other eukaryotic proteins. In contrast, a mutant P68 protein, containing a single amino acid substitution in the invariant lysine in catalytic domain II, was completely inactive. Interestingly, both the mutant and wild-type protein kinases efficiently bound activator dsRNAs despite the fact that only the latter was activated by these RNAs. Finally, the expressed kinase could be isolated from contaminating *E. coli* proteins in an active form by immunoaffinity chromatography with a monoclonal antibody specific for P68.

The P68 protein kinase (referred to as P68 from its molecular weight of 68000 but also referred to by others as DAI, dsl, or PI/eIF2) is a member of the serine/threonine kinase family and one of several genes that can be induced by interferon (Lebleu et al., 1976; Lengyel, 1982; Petska et al., 1982; Samuel, 1979). Two activities are characteristic of P68. The first involves activation by dsRNA or heparin (resulting in the autophosphorylation of the enzyme), and the second is that once phosphorylated, P68 catalyzes the phosphorylation of its natural substrate, the α subunit of eukaryotic protein synthesis initiation factor (eIF-2) (Galabru & Hovanessian, 1987; Hovanessian, 1989). This event prevents the recycling of eIF-2–GDP to eIF-2–GTP by the guanine nucleotide exchange factor eIF-2B, due to the latter being sequestered in an inactive complex with eIF-2–GDP (Konieczny & Safer, 1983; Panniers & Henshaw, 1983; Safer, 1983). Resultant limitations in functional eIF-2 impede the initiation step of protein synthesis by preventing the transfer of initiator Met-tRNA (via the ternary complex eIF-2·GTP·Met-tRNA) to the 40S ribosomal subunit before mRNA is bound (Jagus et al., 1981). Both activation of P68 and phosphorylation of eIF-2 α have been reported to be independent of cAMP and cGMP; both activites, however, are markedly stimulated by the divalent cation manganese and ATP (Hovanessian, 1989).

Activation of P68 and enhanced phosphorylation of eIF-2 α have been documented in virus-infected, interferon-treated cells, suggesting a possible role for P68 in the antiviral response mediated by interferon [for a review, see Hovanessian (1989)]. Indeed, viral-specific RNAs, including those of HIV-1, have the capacity to activate the kinase (Maran & Mathews, 1988; Black et al., 1989; SenGupta et al., 1989; Bischoff & Samuel, 1989; Roy et al., 1991). To counteract the potentially harmful effects caused by this activation, certain viruses have devised strategies to down-regulate the P68 [e.g., see Katze et al. (1987), Lee et al. (1990), Black et al. (1989), Roy et al. (1990), and Mellits et al. (1990)]. Despite these studies, however, little is currently known about the molecular mechanisms involved in activation and repression of P68 or of the potential physiological role of the kinase in the cell.

The gene coding for the human P68 kinase has recently been cloned and sequenced (Meurs et al., 1990). The availability of the cloned gene has enabled us to utilize recombinant expression systems to express P68 for use in structure/function

† This investigation was supported by U.S. Public Health Service Grants AI 22646 and RR 00166 from the National Institutes of Health (M.G.K.). A.G.H. is a research fellow of the Centre Nationale de la Recherche Scientifique, France, of which the Unité de Virologie et Immunologie Cellulaire is a research group (URA 1157).

* Correspondence should be addressed to this author at the Department of Microbiology SC-42, School of Medicine, University of Washington, Seattle, WA 98195
‡ University of Washington.
§ Institut Pasteur.

0006-2960/91/0430-10356$02.50/0