Proc. Fed. Amer. Soc. Exp. Biol. 30, 1093.
Bonhoeffer, F., and Schachman, H. K. (1960), *Biochem. Biophys. Res. Commun. 2*, 366.
Glitz, D. G., and Dekker, C. A. (1963), *Biochemistry 2*, 1185.
Glitz, D. G., and Dekker, C. A. (1964a), *Biochemistry 3*, 1391.
Glitz, D. G., and Dekker, C. A. (1964b), *Biochemistry 3*, 1399.
Grimm, P. W., and Allen, P. J. (1954), *Plant Physiol. 29*, 369.
Harrington, W. F., and Schellman, J. A. (1956), *C. R. Trav. Lab. Carlsberg 30*, 21.
Hirs, C. W., Stein, W. H., and Moore, S. (1954), *J. Biol. Chem. 211*, 941.
Kabat, E. A., and Mayer, M. M. (1961), Experimental Immunochemistry, 2nd ed, Springfield, Ill., C. C Thomas, p 476.
LaBar, F. E. (1965), *Proc. Nat. Acad. Sci. U. S. 54, 31.*
Light, A., and Smith, E. L. (1963), *Proteins 1*, 2.
Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J., (1951), *J. Biol. Chem. 193*, 265.
Moore, S. (1963), *J. Biol. Chem. 238*, 235.
Moore, S., and Stein, W. H. (1954), *J. Biol. Chem. 211*, 907.
Moore, S., and Stein, W. H. (1963), *Methods Enzymol. 6*, 819.
Rushizky, G. W., Greco, A. E., Hartley, R. W., Jr. and Sober, H. A. (1964), *J. Biol. Chem. 239*, 2165.
Sato, K., and Egami, F. (1957), *J. Biochem. (Tokyo) 44*, 753.
Schachman, H. K. (1957), *Methods Enzymol. 4*, 32.
Stark, G. R. (1964), *J. Biol. Chem. 239*, 1411.
Stark, G. R., and Smyth, D. G. (1963), *J. Biol. Chem. 238*, 214.
Takahashi, K. (1965), *J. Biol. Chem. 240*, PC4117.
Takahashi, K., Moore, S., and Stein, W. H. (1967), *J. Biol. Chem. 242*, 4682.
Tanaka, K. (1961), *J. Biochem. (Tokyo) 50*, 62.
Tatarskaya, R. I., Abrosimova-Amel'Yanchik, N. M., Aksel'rod, V. D., Korenyako, A. I., Venkstern, T. V., Mirzabekov, A. D., and Baev, A. A. (1964), *Dokl. Akad. Nauk SSSR 157*, 725.
Teller, D. C. (1965), Ph.D. Thesis, Univ. of California, Berkeley.
Uchida, T. (1965), *J. Biochem. (Tokyo) 57*, 547.
Uchida, T. (1970), *J. Biochem. (Tokyo) 68*, 255.
Uchida, T., and Egami, F. (1966), *in* Procedures in Nucleic Acid Research, Cantoni, G. L., and Davies, D. R., Ed., New York, N. Y., Harper & Row, p 3.
Uchida, T., and Egami, F. (1971), *Enzymes 4*, 209.
Ui, N., and Tarutani, O. (1961), *J. Biochem. (Tokyo) 49*, 9.
Williams, D. E., and Reisfeld, R. A. (1964), *Ann. N. Y. Acad. Sci. 121*, 373.
Yphantis, D. A. (1964), *Biochemistry 3*, 297.

Determination of Nucleotide Sequences at Promoter Regions by the Use of Dinucleotides*

Kathleen M. Downey,† Bonnie S. Jurmark, and Antero G. So‡

ABSTRACT: The nucleotide sequences at the promoter regions and the initiating sequences of RNAs synthesized *in vitro* with T4 or T5 DNAs as templates have been determined by using dinucleotides to initiate RNA synthesis. This method is based on the observations that only dinucleotides that are complementary to the nucleotide sequences at promoter regions, in accordance with Watson–Crick base pairing, can serve as initiators and are incorporated into the 5′ ends of RNA chains. Initiation with dinucleotides was measured by the stimulation of the overall rate of RNA synthesis at low substrate concentrations and by the competition of dinucleotides with [γ-^{32}P]ATP and [γ-^{32}P]GTP for incorporation into the 5′ ends of RNA chains. The initiating dinucleotide sequence of RNA synthesized in the presence of σ factor consists primarily of ApU when either T4 or T5 DNA is used as a template, whereas with core polymerase and T5 DNA as template, the initiating dinucleotide sequence is primarily GpA. The present studies suggest that σ factor recognizes the nucleotide sequence of d(Pu←T←A) at the promotor region of either T4 or T5 DNA, and core polymerase recognizes the sequence d(T←C←T), or a cluster of pyrimidines, in the DNA template. The possible heterogeneity of core polymerase is also discussed.

The regulation of gene expression by positive control elements has now been demonstrated in several microbial systems (Burgess *et al.*, 1969; Travers and Burgess, 1969; Bautz *et al.*, 1969; Losick and Sonenshein, 1969; Travers, 1970). It has been suggested that positive control factors, such as the *E. coli* σ factor, direct the core polymerase to bind to specific promoter sites on the DNA template where RNA chains are initiated. This specificity is reflected in a higher degree of asymmetric transcription in the presence of σ factor (Bautz *et al.*, 1969; Goff and Minkley, 1969; Sugiura *et al.*, 1970).

It is believed that each positive control factor would allow recognition of a unique nucleotide sequence at a specific promoter site or sites on the DNA template. Thus, it is of interest to develop a rapid and reliable method to determine the 5′-terminal nucleotide sequence of *in vitro* synthesized RNA and

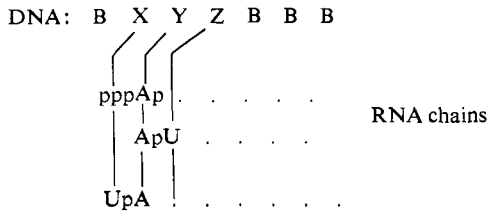
* From the Raymy Haber Karp Hematology Research Laboratories, Department of Medicine and Biochemistry, University of Miami, Miami, Florida 33152. *Received July 21, 1971.* This research was supported by grants from the National Institutes of Health (NIH-AM-09001) and the American Heart Association (69-640).

† To whom to address correspondence.
‡ Established Investigator of the American Heart Association.

the complementary nucleotide sequence at the promoter region. Such a method could serve as an assay for studying gene activation in both prokaryotic and eukaryotic systems, particularly in mammalian systems, where techniques such as hybridization competition are not readily applicable.

In this communication we shall describe a new method to determine the 5′-terminal nucleotide sequence of RNA, as well as the complementary nucleotide sequence in the promoter region of DNA. This method involves the use of dinucleotides to initiate RNA synthesis and is based on our recent findings that: (1) chain initiation is a bimolecular reaction and the rate-limiting step in the synthesis of RNA; (2) complementary dinucleotides, by replacing the initiating nucleotides, greatly stimulate the rate of RNA synthesis at low substrate concentrations; (3) only dinucleotides that are complementary to the initiation sequences of the DNA template can serve as chain initiators, and can be incorporated into 5′ ends of RNA chains; and (4) complementary dinucleotides compete with [γ-^{32}P]ribonucleoside triphosphate for incorporation into the 5′ ends of RNA chains (Downey and So, 1970; So and Downey, 1970; So *et al.*, 1971).

The mapping of the nucleotide sequence at promoter regions of DNA by the use of complementary dinucleotides may be illustrated by the following diagram



Assume that the following were noted. (1) Initiation of RNA chains occurs with pppA, as measured by the incorporation of [γ-^{32}P]ATP. (2) The incorporation of [γ-^{32}P]-ATP is competitively inhibited by the dinucleotides ApU and UpA. (3) UpA and ApU stimulate the overall rate of RNA synthesis at low substrate concentrations. The conclusions drawn would be (1) the nucleotide sequence (Y←Z) at the promoter region is d(T←A) and (2) the nucleotide (X) adjacent and to the left of (Y) is d(A), *i.e.*, the nucleotide sequence at the promoter region is d(A←T←A).

Methods and Materials

RNA polymerase was prepared from *E. coli* cells as by Chamberlin and Berg (1962) and further purified as by So and Downey (1970). σ factor and core polymerase were prepared as by Burgess *et al.* (1969), [γ-^{32}P]ATP and [γ-^{32}P]GTP as by Glynn and Chappell (1964), and T4 DNA as by Bolle *et al.* (1968).

The polymerization assay was carried out as previously reported (Downey and So, 1970). The standard reaction mixture contained 0.08 M Tris-HCl buffer (pH 7.8), 0.2 M KCl, 10 mM $MgCl_2$, 4.8 mM β-mercaptoethanol, 4.6 μg of T4 DNA, 8 μM each of ATP, CTP, GTP, and [^{14}C]UTP, and 10 μg of RNA polymerase (holoenzyme) in a final volume of 0.25 ml. The reaction mixtures were incubated for 10 min at 37°.

The initiation assay was as previously described (Downey and So, 1970) and similar to the polymerization assay except: (1) 16 μM of either [γ-^{32}P]ATP or [γ-^{32}P]GTP was used instead of [^{24}C]UTP and the concentration of each of the remaining three ribonucleoside triphosphates was 0.32 mM; (2) the reaction mixtures were incubated for 5 min at 37°.

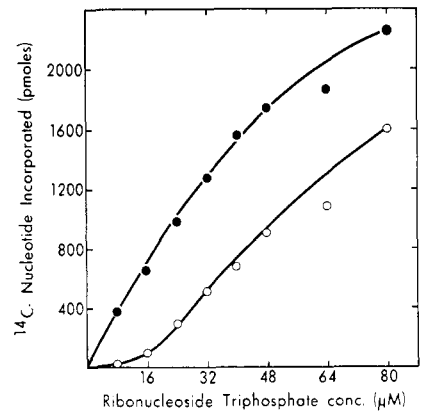


FIGURE 1: Effect of excess ATP on the rate of RNA synthesis at varying substrate concentrations. The reaction mixtures were as described in Methods and Materials except that (1) the concentrations of ribonucleoside triphosphates were as indicated, (2) ● (incubated with additional 0.32 mM ATP), and (3) ○ (no excess ATP was added).

[^{14}C]UTP was purchased from Schwarz BioResearch, Inc. Unlabeled ribonucleoside triphosphates were obtained from either P-L Biochemicals, Inc., or Calbiochem. *E. coli* B cells were obtained from Grain Processing Corp. T5 DNA was purchased from Calbiochem. T4 phage stock was a gift of Dr. B. D. Hall. The dinucleotides were purchased from either Sigma Chemical Co. or Calbiochem and shown to be electrophoretically pure.

Results

Stimulation of RNA Synthesis by High Concentration of Initiating Nucleotide. Using poly[d(A-T)] as a template to study the kinetics of RNA synthesis, we have recently shown that chain initiation, or the formation of the first phosphodiester bond, is a bimolecular reaction and the rate-limiting step in the synthesis of RNA (Downey and So, 1970). Thus the rate of RNA synthesis is extremely sensitive to the concentration of the substrates, the four ribonucleoside triphosphates. At low substrate concentrations, the rate of RNA synthesis is very low. However, if one of the two ribonucleoside triphosphates involved in the formation of the first phosphodiester bond is in excess, both the rate of chain initiation and the overall rate of RNA synthesis are greatly stimulated.

With T4 DNA as template, the RNA synthesized for short intervals with *E. coli* RNA polymerase (core enzyme plus σ factor) in the presence of 0.2 M KCl is initiated primarily with ATP; and the ratio of ATP to GTP initiations under these conditions is approximately 5 to 1 (K. M. Downey and A. G. So, unpublished results). Thus, the rate of RNA synthesis is greatly stimulated by high concentrations of ATP, as shown in Figure 1, where the rate of RNA synthesis is plotted as a function of substrate concentration in the presence and absence of excess ATP. Under these conditions, high concentrations of GTP had little or no stimulatory effect on the rate of RNA synthesis with T4 DNA as a template.

We have also shown that the stimulation effected by an initiating ribonucleoside triphosphate can also be effected by a dinucleotide that is complementary to the initiation sequence on the DNA template (Downey and So, 1970). Thus, with poly[d(A-T)] as a template, the complementary dinucleotides ApU and UpA greatly stimulate the rate of RNA synthesis at low substrate concentrations. No other dinucleotide had any

TABLE I: Stimulation of RNA Synthesis by Dinucleotides with T4 DNA as Template.[a]

Dinucleotides Added	^{14}C Nucleotides Inc (pmoles)
None	10
UpA	96
ApU	62
CpA	22
GpU	17
ApC	16
ApA	16
ApA	15
ApG	15
UpG	15
GpA	14

[a] The reaction mixtures were as described in Methods and Materials except that 0.32 mM dinucleotides were added as indicated.

TABLE II: Effects of Dinucleotides on T4 RNA Chains Initiation with ATP.[a]

Dinucleotides Added	^{32}P Nucleotide Inc (pmoles)	% Inhibn
Control	2.2	
UpA	0.7	68
ApU	1.3	41
CpA	1.4	36
ApA	2.0	9
ApC	2.0	9
UpU	2.0	9
GpU	2.2	0

[a] The reaction mixtures were as described in Methods and Materials except that (1) 80 μM dinucleotides were added as indicated, and (2) 16 μM [γ-^{32}P]ATP and 0.32 mM each of UTP, CTP, and GTP were added.

effect on the rate of poly[r(A-U)] synthesis, a finding demonstrating the specificity of this effect.

Effect of Dinucleotides on RNA Synthesis with T4 DNA as Template. The effects of dinucleotides on the rate of RNA synthesis with T4 DNA as template are shown in Table I. Of the 15 dinucleotides tested (GpG was not tested), only the addition of UpA and ApU resulted in greater than 2-fold stimulation (10-fold and 6-fold, respectively). The other dinucleotides listed in Table I showed between 1.4- and 2-fold stimulation. The dinucleotides not listed in Table I had no effect on RNA synthesis.

Since RNA chains are known to be initiated preferentially with purine ribonucleoside triphosphates (Bremer *et al.*, 1965; Maitra and Hurwitz, 1965), these results would suggest that the most frequent RNA initiation sequence with T4 DNA and *E. coli* holoenzyme is ApU with a smaller amount of ApC and ApA. Thus, the corresponding promoter sequences would be primarily d(T←A), with smaller amounts of d(T←G) and d(T←T). The nucleotide in the promoter region immediately adjacent and to the left of dTMP appears to be predominantly dAMP. A small but significant amount of stimulation was observed with dinucleotides containing guanosine: UpG, GpU, ApG, and GpA. No stimulation was observed with dinucleotides containing only pyrimidines, *i.e.*, UpU, UpC, CpC, and CpU, nor with GpC.

Nucleotide Sequences at Promoter Regions. In order to determine the specificity of the effects of the dinucleotides on the rate of RNA synthesis, to analyze the initiation sequence of the RNA synthesized and to determine the nucleotide sequence at the promoter region, the effects of dinucleotides on the incorporation of [γ-^{32}P]ATP into the 5′ termini of RNA chains have been studied (Table II). The relative degree of inhibition of [γ-^{32}P]ATP incorporation by dinucleotides containing adenosine at the 5′ end (ApN) would yield information on the relative frequency of penultimate RNA nucleotides and the corresponding complementary nucleotides at the promoter region whereas the relative degree of inhibition by NpA dinucleotides would elucidate the relative frequency of nucleotides immediately to the left of the dTMP in the promoter region. The effects of dinucleotides on initiation are rather specific for, except UpU, only adenosine-containing dinucleotides had any inhibitory effect on the incorporation of [γ-^{32}P]ATP. In agreement with the results on the stimulation of the rate of RNA synthesis by dinucleotides, the greatest inhibition of ^{32}P incorporation was observed with UpA, ApU, and CpA. A slight inhibition was found with ApA, ApC, and UpU. These results suggest that the initiation sequence of T4 RNA is ApU and the corresponding complementary nucleotide sequence at the promoter region is d(T←A). Presumably ApU is the 5′-terminal sequence of preearly T4 RNA, since this is the class of RNA that is synthesized during a brief incubation by *E. coli* holoenzyme on a T4 DNA template (Travers, 1970; Bautz *et al.*, 1969; Brody *et al.*, 1970). The results further suggest that the nucleotide sequence recognized by σ factor on the promoter region of the preearly genes of T4 is d(Pu←T←A). That σ factor does indeed recognize the above nucleotide sequence in T4 DNA is supported by the ovservation that only UpA-, ApU-, and CpA-initiated RNA chains are markedly stimulated by σ factor (19-, 12-, and 5-fold, respectively), while chains initiated with UpG, GpU, ApC, and ApA are stimulated less than 2-fold. No stimulation by σ factor was observed with ApG- and GpA-initiated chains (data not shown).

It is not clear why UpU should inhibit [γ-^{32}P]ATP incorporation, even though to a small degree. It is possible that UpU is competing for the template-binding site on the enzyme.

Specificity of Core Polymerase and σ Factor. The exact role of the core polymerase in RNA synthesis has not yet been elucidated. The core polymerase has been shown to transcribe phage DNA more randomly and less asymmetrically than the holoenzyme, although the transcription products do not seem to be completely random (Bautz *et al.*, 1969; Goff and Minkley, 1970; Travers, 1970; Hall *et al.*, 1970). Thus, it is of interest to determine if the core polymerase displays any specificity with respect to RNA initiation.

Because core polymerase transcribes T4 DNA very inefficiently, T5 DNA was used as a template in these experiments. The rate of RNA synthesis with T5 DNA and core polymerase is about 25% of the rate with holoenzyme. The effects of dinucleotides on T5 RNA synthesis catalyzed by core enzyme are shown in Table III. Of the 15 dinucleotides tested, only those listed in Table III stimulated the synthesis of RNA

TABLE III: Effects of Dinucleotides on T5 RNA Synthesis Catalyzed by the Core Enzyme.[a]

Dinucleotides Added	^{14}C Nucleotide Inc (pmoles)	Rel Stimulation (-fold)
Control	5	
GpA	29	5.8
ApG	16	3.2
ApA	13	2.6
GpC	14	2.8
GpU	11	2.2
UpA	10	2.0
UpG	8	1.6

[a] The assay conditions were as described in Table I except that (1) 5 μg of core enzyme was used instead of holoenzyme, and (2) 3.5 μg of T5 DNA was used instead of T4 DNA.

TABLE V: Effects of Dinucleotides on T5 RNA Chain Initiation with ATP.[a]

Dinucleotides Added	^{32}P Nucleotide Inc (pmoles)	% Inhibn
Control	0.53	
CpA	0.23	57
UpA	0.25	53
ApU	0.27	49
ApA	0.45	15
GpA	0.50	6
ApC	0.50	6
UpU	0.51	4
ApG	0.51	4

[a] The assay conditions were as described in Table II except that 3.5 μg of T5 DNA was used instead of T4 DNA.

at low substrate concentrations. GpA and ApG gave the greatest stimulation, 5.3- and 3-fold, respectively. ApA and GpC gave less stimulation, and the addition of GpU and UpA resulted in only a slight stimulation. These results suggest that the core enzyme does exercise some selectivity in initiating RNA synthesis, preferentially utilizing dinucleotides containing guanosine in the 5′ position and the purine–purine dinucleotides.

When holoenzyme is used with T5 DNA as a template, the dinucleotides causing stimulation are shown in Table IV. Similar to the results obtained with T4 DNA and holoenzyme, UpA and ApU gave the greatest stimulation, followed by CpA and ApA, suggesting that σ factor recognizes the same nucleotide sequence(s) in different DNA templates. In contrast with the results obtained with T4, however, the dinucleotides GpA and ApG also stimulated T5 RNA synthesis significantly.

The stimulation of RNA synthesis by the dinucleotides ApG and GpA could indicate that these dinucleotides are functioning as either ATP or GTP initiators. This question is resolved by studying the effects of these dinucleotides on the incorporation of [γ-^{32}P]ATP (Table V) and [γ-^{32}P]GTP (Table VI).

TABLE IV: Stimulation of T5 RNA Synthesis by Dinucleotides with Holoenzyme.[a]

Dinucleotides Added	^{14}C Nucleotide Inc (pmoles)	Rel Stimulation (-fold)
None	8	
UpA	51	6.4
ApU	48	6.0
GpA	32	4.0
CpA	28	3.5
ApG	27	3.4
ApA	25	3.1
GpU	18	2.3
GpC	16	2.0
UpG	15	1.9
CpG	11	1.4

[a] The assay conditions were as described in Table I except that 3.5 μg of T5 DNA was used instead of T4 DNA.

ApG and GpA markedly inhibit the rate of incorporation of [γ-^{32}P]GTP while the incorporation of [γ-^{32}P]ATP is only very slightly inhibited. The rate of incorporation of [γ-^{32}P]-GTP is also inhibited by UpG, CpG, GpU, and GpC whereas [γ-^{32}P]ATP incorporation is inhibited CpA, UpA, ApU, and ApA.

These results on the inhibition of GTP initiation, coupled with the findings on dinucleotide-stimulated RNA synthesis catalyzed by core polymerase, strongly suggest that the initiating sequences of RNAs synthesized with core polymerase are primarily GpA and ApA, with significant amounts of GpU and GpC. The results also indicate that the core polymerase has a special affinity for clusters of pyrimidines in the DNA template, specifically d(T←C←T).

It has been suggested that σ factor confers specificity on the core polymerase, enabling it to transcribe specific genes on the DNA template. We have, therefore, investigated the effects of σ factor on core-catalyzed RNA initiation. The effects of σ factor on dinucleotide-stimulated T5 RNA synthesis are shown in Table VII. Little or no stimulation by σ factor was observed with the GpN dinucleotides (GpA, GpU, GpC) or with the purine–purine dinucleotides (ApA, ApG, GpA). However, the presence of σ factor did not inhibit the synthesis of RNA initiated with these dinucleotides. Thus, it would appear that the binding of σ factor to the core polymerase

TABLE VI: Effects of Dinucleotides on T5 RNA Chain Initiation with GTP.

Dinucleotides Added	^{32}P Nucleotide Inc (pmole)	% Inhibn
Control	1.00	
UpG	0.61	39
CpG	0.65	35
GpA	0.69	31
ApG	0.70	30
GpU	0.79	21
GpC	0.82	18

[a] The assay conditions were as described in Table V except that 16 μM [γ-^{32}P]GTP was used instead of [γ-^{32}P]ATP.

TABLE VII: Effects of σ Factor on Dinucleotide-Stimulated T5 RNA Synthesis.[a]

Dinucleotides Added	^{14}C Nucleotide Inc (pmoles)		
	Core Enzyme	Core Enzyme + σ Factor	Rel Stimulation by σ (-fold)
Low NTP	5	8	1.6
High NTP	9	45	5.0
ApU	6	48	9.6
UpA	10	51	5.1
CpA	5	28	5.6
ApA	13	25	1.9
ApG	16	27	1.7
GpU	11	18	1.6
GpA	29	32	1.1
UpG	8	17	2.1

[a] The assay conditions were as described in Table III except that 3.8 μg of σ factor was added as indicated. The concentrations of ATP, GTP, CTP, and [^{14}C]UTP at low NTP were 16 μM and at high NTP were 0.32 mM.

results in enzyme recognition of new promoter sequences, while those sequences recognized by the core enzyme continue to be recognized by the holoenzyme.

Discussion

The present studies clearly demonstrate the feasibility and ease of using dinucleotides to map nucleotide sequences at promoter regions and to determine RNA initiation sequences in an *in vitro* system. This technique should be applicable to studies on the mechanism of gene activation in eukaryotic systems as well as in prokaryotic systems. Determination of nucleotide sequences around the initiating 5′ termini by the use of dinucleotides greatly amplifies the specificity of studies on gene activation, as compared to determination of only the 5′ termini. The present technique may be extended to determine longer initiating nucleotide sequences by the use of oligonucleotides of greater length, being limited only by the availability of oligonucleotides of known sequence.

The rationale for using dinucleotides to determine the initiation sequences of RNAs is based on the following considerations: (1) dinucleotides can only be incorporated into the 5′ ends of RNA chains and not internally, and (2) only dinucleotides that are complementary to the nucleotide sequences at promoter regions of the DNA template, in accordance with the rules of Watson–Crick base pairing, can serve as chain initiators and be incorporated into the 5′ ends of RNA chains (Downey and So, 1970; So and Downey, 1970; So *et al.*, 1971).

In the present studies, RNA initiation sequences have been determined with two different templates, T4 and T5 DNA, both of which are normal templates for the *E. coli* enzyme. With T4 DNA very little initiation takes place with core enzyme, and initiation appears to occur predominantly with ATP. With T5 DNA as template, both holoenzyme and core polymerase can initiate RNA synthesis, and initiation occurs with both ATP and GTP. Thus, by comparing results obtained with both templates, it has been possible to demonstrate: (1) that core polymerase shows some specificity in initiating RNA chains, preferentially recognizing the d(T←C←T) sequence on the template and initiating RNA chains predominantly with GTP, (2) that σ factor causes recognition of the d(Pu←T←A) sequence on both T4 and T5 DNA, and (3) that the addition of σ factor to core polymerase causes recognition of new promoter sequences, in addition to those recognized by core alone.

At present, the length of the nucleotide sequence required to confer specificity at a promoter site is not known. Presumably the secondary structure of the DNA plays an important role, since the specificity of RNA initiation is lost when the template is denatured (Richardson, 1969). However, RNA synthesis is still preferentially initiated with purine nucleotides, even with denatured DNA (Maitra and Hurwitz, 1965).

It has been suggested that clusters of pyrimidines in the DNA template are the initiation signals for RNA transcription (Szybalski *et al.*, 1966). The present results indicate that the core polymerase appears to have an affinity for pyrimidine clusters in the template (*e.g.*, d(T←C←T) or d(T←T), however, the holoenzyme seems to recognize preferentially an alternating pyrimidine–purine signal (*e.g.*, d(T←A←T)).

Nearest-neighbor analysis of nucleotides adjacent to the 5′-adenylate termini of T4 RNA has shown that the penultimate nucleotides are predominantly pyrimidines (38% cytidine and 32% uridine) (Maitra *et al.*, 1967). Our studies with dinucleotide-initiated T4 RNA synthesis have also shown that the predominant penultimate nucleotides are pyrimidines, however, UMP is by far the most common. This is in agreement with the results of Bautz *et al.* (1970) who showed that ApU was the most frequent initiation sequence of T4 RNA using a pyrophosphate-exchange assay.

The present observation that RNA chains initiated with ApU are markedly stimulated by σ factor while little or no stimulation was observed with GpA- and other G-initiated RNA chains is consistent with the observation of Goff and Minkley (1969) that with T7 DNA, adenylate 5′ termini and not guanylate 5′ termini, are stimulated by σ factor. Our finding that the initiating sequence of RNA catalyzed by core polymerase consists largely of purine–purine sequences (GpA and ApA) might explain the observation of Goff and Minkley (1970) that 5′ termini synthesized with core polymerase are resistant to degradation by pancreatic RNase.

Although the present studies indicate that core polymerase does show specificity in initiating RNA chains, it is not known whether this core-catalyzed RNA has any functional role *in vivo*.

It is interesting to note that the addition of excess σ factor to core polymerase, although increasing ApU-initiated RNA synthesis, does not affect the rate of core-specific initiation with GpA. One possible explanation for this result is that the core polymerase is heterogeneous and consists of two species of core molecules, one species being inactive and requiring σ factor to recognize the d(T←A←T) sequence, and the other species being active, recognizing the initiation sequences we have designated as core specific (*e.g.*, d(T←C←T)), and having little or no affinity for σ factor. This hypothesis is consistent with the observation that the addition of σ factor to core enzyme results in a twofold increase in the amount of T7 DNA retained on a nitrocellulose filter (Hinkle and Chamberlin, 1970). Also consistent with this hypothesis is the observation of Zillig *et al.* (1970) that holoenzyme contains only 50% of the stoichiometrically saturating amount of σ factor.

If the core polymerase is heterogeneous, the physical differ-

ences between the two molecules must be small, since no evidence of heterogeneity can be detected by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate.

References

Bautz, E. K. F., Bautz, F. A., and Dunn, J. J. (1969), *Nature (London) 223*, 1022.

Bautz, E. K. F., Dunn, J. J., Bautz, F. A., Schmidt, D. A., and Mazaitis, A. J. (1970), *in* Lepetit Colloquim on RNA Polymerase and Transcription, Silvester, L., Ed., Amsterdam, North-Holland, p 90.

Bolle, A., Epstein, R. H., Salser, W., and Geiduschek, E. P. (1968), *J. Mol. Biol. 31*, 325.

Bremer, H., Konrad, M. W., Gaines, K., and Stent, G. S. (1965), *J. Mol. Biol. 13*, 540.

Brody, E., Sederoff, R., Bolle, A., and Epstein, R. (1970), *Cold Spring Harbor Symp. Quant. Biol. 35*, 203.

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

Chamberlin, M., and Berg, P. (1962), *Proc. Nat. Acad. Sci. U. S. 48*, 81.

Downey, K. M., and So, A. G. (1970), *Biochemistry 9*, 2520.

Glynn, I. M., and Chappell, J. B. (1964), *Biochem. J. 90*, 147.

Goff, C. G., and Minkley, E. G. (1970), *in* Lepetit Colloquium on RNA Polymerase and Transcription, Silvester, L., Ed., Amsterdam, North-Holland, p 124.

Hall, B. D., Fields, K., and Hager, G. (1970), *in* Lepetit Colloquium on RNA Polymerase and Transcription, Silvester, L., Ed., Amsterdam, North-Holland, p 148.

Hinkle, D., and Chamberlin, M. J. (1970), *Cold Spring Harbor Symp. Quant. Biol. 35*, 65.

Losick, R.,and Sonenshein, A. L. (1969), *Nature (London) 224*, 35.

Maitra, U., and Hurwitz, J. (1965), *Proc. Nat. Acad. Sci. U. S. 54*, 815.

Maitra, U., Nakata, Y., and Hurwitz, J. (1967), *J. Biol. Chem. 242*, 4908.

Richardson, J. P. (1969), *Progr. Nucl. Acid Res. Mol. Biol. 9*, 65.

So, A. G., and Downey, K. M. (1970), *Biochemistry 9*, 4788.

So, A. G., Downey, K. M., and Jurmark, B. S. (1971), *in* Miami Winter Symposia, Vol. 2, Ribbons, D. W., Woessner, J. F., and Schultz, J., Ed., Amsterdam, North-Holland, p 221.

Sugiura, M., Okamoto, T., and Takanami, M. (1970), *Nature (London) 225*, 598.

Szybalski, W., Kubinski, H., and Sheldrick, P. (1966), *Cold Spring Harbor Symp. Quant. Biol. 31, 123.*

Travers, A. A. (1970), *Nature (London) 225*, 1009.

Travers, A. A., and Burgess, R. R. (1969), *Nature (London) 222*, 537.

Zillig, W., Palm, P., Sethi, V. S., Zechel, K., Rabussay, D., Heil, A., Seifart, W., and Schachner, M. (1970), *Cold Spring Harbor Symp. Quant. Biol., 35*, 47.

F_1 Sex Factor of *Escherichia coli*. Size and Purification in the Form of a Strand-Specific Relaxation Complex of Supercoiled Deoxyribonucleic Acid and Protein*

Bruce C. Kline† and Donald R. Helinski‡

ABSTRACT: Supercoiled F_1 sex factor DNA complexed with cellular material was isolated from logarithmically growing *Escherichia coli* cells. The sedimentation coefficient of the complex is approximately 84 S whereas noncomplexed, supercoiled F_1 DNA is 80 S. Treatment of a mixture of complexed and noncomplexed supercoiled DNA with proteases, alkali, or sodium dodecyl sulfate results in a 50% conversion of the complexed DNA to the open-circular form of DNA with no significant effect on the noncomplexed DNA. The open-circular form resulting from this conversion possesses a nick or gap specifically in the denser of the two strands of F_1 DNA separated by equilibrium centrifugation in a CsCl gradient containing poly(U,G). The average contour length of the open-circular form of F_1 DNA was determined to be 31.7 ± 2.1 μ corresponding to a molecular weight of 60.9 ± 4.1 × 10^6.

Several supercoiled extrachromosomal elements have been isolated from *Escherichia coli* as DNA–protein complexes, for example, colicinogenic factors *Col*E_1 (K30) (Clewell and Helinski, 1969), *Col*E_2 (P9) (Clewell and Helinski, 1970a; Blair *et al.*, 1971), *Col*E_3 (CA38) (Clewell and Helinski, 1970a), and the colicinogenic factor–sex factor, *Col*I_b (P9) (Clewell and Helinski, 1970b). A characteristic property of these supercoiled DNA–protein complexes, designated relaxation complexes, is their conversion to the open-circular DNA form possessing a nick or gap in one of the two DNA strands after

* From the Department of Biology, University of California, San Diego, La Jolla, California 92037. *Received June 28, 1971.* Supported by U. S. Public Health Service Research Grant AI-07194 and National Science Foundation Grant GB-6297. B. C. K. gratefully acknowledges a U. S. Public Health Service, National Cancer Institute postdoctoral fellowship (I-F02-CA43, 050). D. R. H. is a U. S. Public Health Service Research Career Development awardee (K04-6M07821).

† Present address: Department of Biochemistry, University of Tennessee, Knoxville, Tenn. 37916.

‡ To whom correspondence should be addressed.