*THE ENZYMATIC SYNTHESIS OF RNA: NEAREST-NEIGHBOR BASE FREQUENCIES**

By Samuel B. Weiss and Tokumasa Nakamoto

ARGONNE CANCER RESEARCH HOSPITAL† AND DEPARTMENT OF BIOCHEMISTRY, UNIVERSITY OF CHICAGO

Communicated by Charles Huggins, July 31, 1961

RNA[1] polymerase activity, first detected in mammalian cell-free preparations,[2, 3] has now been described in a number of bacterial extracts and plant tissues.[4–7] The reaction catalyzed by this enzyme shows an absolute requirement for all four ribonucleoside triphosphates and DNA and may be represented by the following equation:

$$\begin{matrix} n\text{ATP} \\ + \\ n\text{GTP} \\ + \\ n\text{CTP} \\ + \\ n\text{UTP} \end{matrix} \xrightleftharpoons{\text{DNA}} \left[\begin{matrix} \text{Ap} \\ \text{Gp} \\ \text{Cp} \\ \text{Up} \end{matrix}\right]_n 4n\text{PP}.$$

The requirement for a deoxypolynucleotide primer raised the question of the role played by this polymer in RNA synthesis. A recent report from this laboratory demonstrated that the base composition of RNA, prepared with a partially purified extract from *Micrococcus lysodeikticus*, was similar to and dependent upon the composition of the primer.[8] Similar results have been obtained by Stevens[9] and Furth *et al.*[10, 11] with extracts prepared from *Escherichia coli*.

Cellular RNA similar in base composition to DNA has been identified in various organisms by a number of investigators.[12, 13] Recently, Hall and Spiegelman have shown that the RNA which appears after T2 phage infection of *E. coli* forms complexes with denatured phage DNA.[14] Such complexes are believed to occur only between strands of complementary base sequences and point out the complementary relationship between the deoxy- and ribopolynucleotides. The existence of similar DNA and RNA molecules in nature has added impetus to the idea that genetic information is transmitted to other cellular sites by so-called "messenger-RNA."[15, 16] The discovery of a DNA-dependent RNA polymerase has stimulated speculation that this enzyme may be responsible for the synthesis of this specific-RNA. So far, no direct experimental proof has been offered to support these speculations.

Weiss and Nakamoto have presented preliminary evidence indicating that the relative position of ribonucleotides in enzymatically prepared RNA is determined by the primer used for the reaction.[8] This information, along with the similar base composition found, suggests, but does not prove, that synthesis occurs with replication of the nucleotide sequences of the primer. Absolute and unequivocal proof of replication can only be supplied by determining the sequential arrangement of bases in both primer and RNA product. In the absence of methods adequate to perform such analysis, a complete nearest-neighbor frequency study for the 16 possible nucleotide pairs in RNA has been determined for RNA's prepared with

different primers. In view of the elegant studies reported recently by Josse, Kaiser, and Kornberg,[17] it is possible to compare our nearest-neighbor pairs with the results published for DNA's that have been similarly characterized. We have also carried out a series of CsCl density-gradient experiments designed to determine the complementary relationship of the RNA formed during reaction with respect to the primer employed, and these results are reported in the following paper.

Materials and Methods.—ATP^{32}, CTP^{32}, and UTP^{32} were prepared as previously described.[3] GTP^{32} was prepared by the method of Tener.[18] Each of these labeled substrates contained P^{32} in the innermost phosphate only. RNA polymerase was obtained from extracts of *M. lysodeikticus*, and preparations ranging from 100- to 200-fold purification were used. *E. coli* DNA was a gift from Dr. A. Rich; T2 DNA was a gift from Dr. L. Grossman; calf thymus DNA was purchased from the Sigma Chemical Company, 3500 DeKalb Street, St. Louis 18, Missouri; and *M. lysodeikticus* DNA was prepared from lysozyme-treated cells as described by Lehman and co-workers.[19]

The reaction mixture for the preparation of P^{32}-labeled RNA contained 3 μmole of $MnCl_2$, 100 μmole Tris-PO_4 buffer of pH 7.4, 0.10 mg of a DNA primer, 1 μmole each of ATP, UTP, CTP, and GTP, only one of which was radioactive (specific activity of 1 to 10 $\times$ 10^6 cpm per μmole), and a purified extract from *M. lysodeikticus* in a total volume of 1.0 ml. Sufficient enzyme was used so that a minimum of 100,000 cpm was incorporated into RNA. For each DNA used, four similar reactions were prepared except that each reaction contained a different labeled substrate. After 30 min at 25°, the reaction was stopped with cold 5 per cent TCA and iced, and the precipitate was collected. The acid-insoluble material was washed three times by dissolving in 0.10 *M* NaOH and reprecipitating with TCA, and once again with ethanol-ether (3:1). One mg of carrier RNA was added, and the residue was extracted twice in 1 ml of 10 per cent NaCl, pH 8, for 30 min at 100°. Two volumes of cold ethanol were added to the combined extracts, and the precipitate was collected after standing at 0° overnight. The nucleic acid fraction was dissolved in 1 ml of water, 0.10 ml of a 20 per cent potassium acetate solution (adjusted to pH 5) was added, and reprecipitation was with ethanol. This procedure was repeated until the ethanol supernatant gave a constant amount of radioactivity. The precipitate was dried under vacuum for 30 min, hydrolyzed in 0.3 *M* KOH for 18 hr at 37°, and acidified with $HClO_4$, and the acid-soluble fraction was subjected to paper electrophoresis as described elsewhere.[3] The 2′-, 3′-nucleoside monophosphates were identified under ultraviolet light, eluted, and assayed for radioactivity.

Results.—During alkaline hydrolysis of RNA, the phosphate esterified to the 5′ position of one ribonucleotide is transferred to the 2′,3′ position of the adjacent nucleotide. P^{32}-RNA, enzymatically synthesized with one ribonucleoside 5′ (pyro)-triphosphate labeled in the ester phosphate only, gives rise to four labeled 2′,3′-nucleoside monophosphates when subjected to this type of hydrolysis. Quantitative determination of the amount of isotope contained in each mononucleotide may be used as a measure of the relative frequency with which the labeled nucleotide occurs next to a given neighbor in the RNA chain. In the synthesis of RNA, if four labeled nucleotides are used separately, in four different reactions, the isotope content of the hydrolyzed mononucleotides can be used to determine the frequencies

TABLE 1

P^{32} CONTENT OF MONONUCLEOTIDES ISOLATED FROM ENZYMATICALLY SYNTHESIZED P^{32}-RNA PRIMED BY *M. lysodeikticus* DNA

Isolated 2',3'-ribonucleotide	Labeled Triphosphate: Reaction No. 1 ATP^{32} Sequence	Cpm	Fraction	Reaction No. 2 UTP^{32} Sequence	Cpm	Fraction	Reaction No. 3 GTP^{32} Sequence	Cpm	Fraction	Reaction No. 4 CTP^{32} Sequence	Cpm	Fraction
Up	UpA	960	0.065	UpU	947	0.108	UpG	3,834	0.151	UpC	8,124	0.185
Ap	ApA	1,612	0.109	ApU	1,411	0.161	ApG	3,237	0.128	ApC	6,667	0.152
Cp	CpA	5,895	0.397	CpU	3,249	0.371	CpG	10,335	0.406	CpC	13,548	0.310
Gp	GpA	6,360	0.429	GpU	3,145	0.360	GpG	8,041	0.315	GpC	15,447	0.353
Sum	..	14,827	1.000	..	8,752	1.000	..	25,447	1.000	..	43,786	1.000

TABLE 2

P^{32} CONTENT OF MONONUCLEOTIDES ISOLATED FROM ENZYMATICALLY SYNTHESIZED P^{32}-RNA PRIMED BY T2 DNA

Isolated 2',3'-ribonucleotide	Labeled Triphosphate: Reaction No. 1 ATP^{32} Sequence	Cpm	Fraction	Reaction No. 2 UTP^{32} Sequence	Cpm	Fraction	Reaction No. 3 GTP^{32} Sequence	Cpm	Fraction	Reaction No. 4 CTP^{32} Sequence	Cpm	Fraction
Up	UpA	3,735	0.258	UpU	3,724	0.340	UpG	6,001	0.350	UpC	8,576	0.343
Ap	ApA	5,400	0.373	ApU	3,544	0.324	ApG	5,098	0.298	ApC	6,818	0.273
Cp	CpA	2,760	0.190	CpU	1,878	0.171	CpG	2,856	0.166	CpC	4,336	0.173
Gp	GpA	2,580	0.179	GpU	1,811	0.165	GpG	3,170	0.186	GpC	5,275	0.211
Sum	..	14,475	1.000	..	10,957	1.000	..	17,125	1.000	..	25,005	1.000

of the 16 possible nearest-neighbor base sequences. For a detailed description of nearest-neighbor sequence determination, the reader is referred to Josse and coworkers.[17] The method used here for calculating these values is essentially as reported by these investigators and may be simply described as follows:

Fractional base value × *base-incorporating factor* = *relative frequency.*

For a given reaction, the isotope content found for each mononucleotide was converted to the decimal fraction of the total to give the fractional base values (Table 1). To obtain nearest-neighbor frequencies, the fractional values must be multiplied by a factor which represents the relative amount of each ribonucleotide incorporated into RNA. Since the relative amount of each ribonucleotide incorporated is determined by the base composition of the primer,[8] the molar proportion of adenine, thymine, guanine, and cytosine of the particular DNA employed in each reaction were used as the base-incorporating factors. The DNA base compositions given in Table V of the paper by Josse, Kaiser, and Kornberg[17] were used, since it was felt that these values might be more accurate than previously reported determinations of the chemical composition for deoxyribonucleic acid.

DNA from *M. lysodeikticus* contains a relatively high guanine and cytosine content while T2 phage DNA is rich in adenine

and thymine. Table 1 and Table 2 show the counts and fractional values determined for the various 2′,3′-mononucleotides formed by alkaline hydrolysis of P^{32}-RNA when synthesized in the presence of these distinctly different primers. When the fractional values are multiplied by the appropriate base-incorporating factors (0.147, 0.145, 0.354, and 0.354 for *M. lysodeikticus* DNA and 0.319, 0.318, 0.184, and 0.179 for T2 DNA in the reactions containing ATP^{32}, UTP^{32}, GTP^{32}, and CTP^{32}, respectively), the nearest-neighbor frequencies shown in Table 3 and Table 4 are obtained.

TABLE 3

COMPARISON OF NEAREST-NEIGHBOR FREQUENCIES OF RNA WITH DNA FOR PRIMER *M. lysodeikticus* DNA

Reaction no.	Labeled triphosphate	Isolated 2′,3′-nucleoside Monophosphate							
		Up (Tp)*		Ap		Cp		Gp	
1	ATP^{32}	UpA	(TpA)	ApA		CpA		GpA	
		0.010	(0.011)	0.016	(0.019)	0.058	(0.052)	0.063	(0.065)
2	UTP^{32}	UpU	(TpU)	ApU	(ApT)	CpU	(CpT)	GpU	(GpT)
	($dTTP^{32}$)	0.016	(0.017)	0.023	(0.022)	0.054	(0.050)	0.052	(0.056)
3	GTP^{32}	UpG	(TpG)	ApG		CpG		GpG	
		0.053	(0.054)	0.045	(0.049)	0.143	(0.139)	0.112	(0.112)
4	CTP^{32}	UpC	(TpC)	ApC		CpC		GpC	
		0.065	(0.063)	0.054	(0.054)	0.110	(0.113)	0.125	(0.121)

* The values in parentheses represent the nucleotide pairs and nearest-neighbor frequencies reported by Josse *et al*[17] for *M. lysodeikticus* DNA.

TABLE 4

COMPARISON OF NEAREST-NEIGHBOR FREQUENCIES OF RNA WITH DNA FOR PRIMER T2 DNA

Reaction no.	Labeled triphosphate	Isolated 2′,3′-nucleoside Monophosphates							
		Up (Tp)*		Ap		Cp		Gp	
1	ATP^{32}	UpA	(TpA)	ApA		CpA		GpA	
		0.082	(0.089)	0.119	(0.111)	0.061	(0.061)	0.057	(0.059)
2	UTP^{32}	UpU	(TpU)	ApU	(ApT)	CpU	(CpT)	GpU	(GpT)
	($dTTP^{32}$)	0.108	(0.106)	0.103	(0.104)	0.054	(0.054)	0.052	(0.051)
3	GTP^{32}	UpG	(TpG)	ApG		CpG		GpG	
		0.064	(0.063)	0.055	(0.057)	0.031	(0.030)	0.034	(0.036)
4	CTP^{32}	UpC	(TpC)	ApC		CpC		GpC	
		0.061	(0.057)	0.049	(0.048)	0.031	(0.034)	0.038	(0.040)

* The values in parentheses represent the nucleotide pairs and nearest-neighbor frequencies reported by Josse *et al.*[17] for T2 DNA.

A comparison of the figures of similar base-pairs found in RNA with those reported for DNA (values in parentheses) shows that the frequency of the 16 nucleotide pairs for primer and product is remarkably close. In all instances where the base-pair frequencies are relatively high or low for a given DNA, the same variation is detected in the base-pairs for the synthesized RNA.

E. coli and calf thymus DNA have base compositions which are somewhat alike. The nearest neighbor frequencies for both these DNA's are also fairly similar except in one or two instances, e.g., the base-pairs CpG and GpC. When the nucleotide pairs of the RNA's formed with these two primers are examined, we find that the frequencies are in good agreement with the values reported for DNA (Table 5). In those few instances where there are marked differences in the base-pair frequencies of the primers, there are also similar differences in the RNA synthesized. Thus, when differences in the nearest-neighbor sequences of two primers exist,

TABLE 5

COMPARISON OF NEAREST-NEIGHBOR FREQUENCIES FOR RNA AND DNA

Nearest-neighbor sequence*	Calf Thymus		*E. coli*	
	Primer DNA†	Synthesized RNA	Primer DNA†	Synthesized RNA
ApA, UpU	0.089, 0.087	0.093, 0.082	0.071, 0.076	0.075, 0.079
CpA, UpG	0.080, 0.076	0.082, 0.070	0.071, 0.071	0.071, 0.068
GpA, UpC	0.064, 0.067	0.062, 0.061	0.055, 0.056	0.059, 0.056
CpU, ApG	0.067, 0.072	0.072, 0.067	0.055, 0.055	0.051, 0.051
GpU, ApC	0.056, 0.052	0.058, 0.053	0.055, 0.054	0.054, 0.058
GpG, CpC	0.050, 0.054	0.061, 0.055	0.056, 0.056	0.057, 0.058
UpA	0.053	0.051	0.051	0.044
ApU	0.073	0.068	0.068	0.070
CpG	0.016	0.016	0.067	0.073
GpC	0.044	0.048	0.083	0.077

* The nucleotide pairs listed are for RNA and are the same for DNA except that T (thymine) would be substituted for U (uridine).

† The nearest-neighbor frequencies for calf thymus and *E. coli* DNA have been reported by Josse *et al.*[17]

the corresponding differences in the RNA products can be readily detected.

Discussion and Summary.—A thorough understanding of the enzymatic mechanism for the synthesis of RNA by DNA-dependent enzymes is important for obvious reasons. The implications of this reaction as a possible step in transmitting genetic information has already been mentioned. The results tabulated in this paper, together with the similar base compositions of DNA and RNA previously determined, provide further evidence for the sequential assembly of ribonucleotides along a DNA template. The data are consistent with the idea that the arrangement of bases along such a template would occur by hydrogen-bonding of the appropriate base-pairs.

RNA polymerase can be primed by heat-denatured DNA as well as native DNA.[20] No definitive answer can be given to the problem whether one or both strands of DNA determine the base sequence and composition of the enzymatically synthesized RNA. In theory, according to the Watson and Crick base-pairing hypothesis,[21] opposing strands of DNA need not have the same base composition, but they may do so on a purely statistical basis. The similar base ratios found for RNA and its primer DNA and the similarity in nearest-neighbor frequencies are consistent with two possibilities: (1) the strands are different and both prime, or (2) the strands are alike in compositions and base sequence and one or both may prime. In the latter case, the results offer no information on the mechanism of priming.

Although the nearest-neighbor frequencies determined in these experiments suggest that the sequential arrangement of nucleotides in RNA is probably the same as that found in the primer, the above results by themselves cannot be considered unequivocal evidence for faithful replication of the base sequences found in DNA. For this reason, we have carried out CsCl density-gradient studies which will be reported in the subsequent paper[22] and which provide more conclusive evidence that the RNA strands synthesized are complementary in base sequence to the deoxypolynucleotide primer. The term complementary is defined as a specific fit of portions of primer and product by appropriate base interactions.

* This investigation was supported by funds from the Argonne Cancer Research Hospital and from the Joseph and Helen Regenstein Foundation.

† Operated by the University of Chicago for the U.S. Atomic Energy Commission.

[1] The abbreviations used in this report are: RNA and DNA for ribo- and deoxyribonucleic

acid; CTP^{32}, ATP^{32}, UTP^{32}, and GTP^{32} for the triphosphates of cytidine, adenosine, uridine, and guanosine labeled with P^{32} in the ester phosphate; Cp, Ap, Up, and Gp for the 2′,3′-nucleoside monophosphates of the bases listed above; PP for inorganic pyrophosphate; Tris, tris-(hydroxymethyl)-aminomethane; TCA, trichloroacetic acid; cpm, counts per minute.

[2] Weiss, S. B., and L. Gladstone, *J. Am. Chem. Soc.*, **81**, 4118 (1959).

[3] Weiss, S. B., these PROCEEDINGS, **46**, 1020 (1960).

[4] Stevens, A., *Biochem. and Biophys. Research Communs.*, **3**, 92 (1960).

[5] Hurwitz, J., A. Bresler, and R. Diringer, *Biochem. and Biophys. Research Communs.*, **3**, 15 (1960).

[6] Weiss, S. B., and T. Nakamoto, *J. Biol. Chem.*, **236**, PC18 (1961).

[7] Huang, R. C., N. Maheshwari, and J. Bonner, *Biochem. and Biophys. Research Communs.*, **3**, 15 (1960).

[8] Weiss, S. B., and T. Nakamoto, these PROCEEDINGS, **47**, 694 (1961).

[9] Stevens, A., *J. Biol. Chem.*, **236**, PC43 (1961).

[10] Furth, J. J., J. Hurwitz, and M. Goldmann, *Biochem. and Biophys. Research Communs.*, **4**, 362 (1961).

[11] *Ibid.*, 431 (1961).

[12] Volkin, E., and L. Astrachan, *Virology*, **2**, 149 (1956).

[13] Yčas, M., and W. S. Vincent, these PROCEEDINGS, **46**, 804 (1960).

[14] Hall, B. D., and S. Spiegelman, these PROCEEDINGS, **47**, 137 (1961).

[15] Brenner, S., F. Jacob, and M. Meselson, *Nature*, **190**, 576 (1961).

[16] Gros, F., H. Hiatt, W. Gilbert, C. G. Kurland, R. W. Rissebrough, and J. D. Watson, *Nature*, **190**, 581 (1961).

[17] Josse, J., A. D. Kaiser, and A. Kornberg, *J. Biol. Chem.*, **236**, 864 (1961).

[18] Tener, G. M., *J. Am. Chem. Soc.*, **83**, 159 (1961).

[19] Lehman, I. R., S. B. Zimmerman, J. Adler, M. J. Bessman, E. S. Simms, and A. Kornberg, these PROCEEDINGS, **44**, 1191 (1958).

[20] Weiss, S. B., *Federation Proc.* (in press).

[21] Watson, J. D., and F. H. C. Crick, *Nature*, **171**, 737, 964 (1953).

[22] Geiduschek, E. P., T. Nakamoto, and S. B. Weiss, these PROCEEDINGS, **47**, 1405 (1961).

---

*THE ENZYMATIC SYNTHESIS OF RNA: COMPLEMENTARY INTERACTION WITH DNA**

BY E. PETER GEIDUSCHEK, TOKUMASA NAKAMOTO, AND SAMUEL B. WEISS

COMMITTEE ON BIOPHYSICS, ARGONNE CANCER RESEARCH HOSPITAL† AND DEPARTMENT OF BIOCHEMISTRY, UNIVERSITY OF CHICAGO

Communicated by Charles Huggins, July 31, 1961

The chemical properties of the polynucleotides synthesized by the DNA-primed RNA polymerase have been described in two previous communications.[1, 2] It has been shown that the average composition and the average nearest-neighbor frequencies of enzymatically synthesized RNA and DNA primer are identical. In view of this information and of the results communicated here, we have named the product of this reaction "Complementary RNA" (C-RNA).[3] It is our working hypothesis that C-RNA is structurally and functionally similar to "messenger" RNA,[9, 10] and that the biosynthesis of C-RNA is an initial step in that sequence of reactions in which DNA asserts the organism's potentialities for protein synthesis. Consequently, it is of considerable interest to discover whether higher-order DNA nucleotide *sequence* is translated with any degree of fidelity in RNA and to describe