Determination of the Molecular Weight in Nucleic Acids by End-Phosphate Analysis†

Libero Clerici, Enrico Sabbioni, Francesco Campagnari,* Silvio Spadari, and Francesco Girardi

ABSTRACT: The differential measurements of total and terminal phosphorus in nucleic acids were brought to very high sensitivity and resolution by using alkaline phosphatase from calf thymus to reveal the monoesterified phosphate and neutron activation analysis to label the two phosphorus fractions with ^{32}P. Chain lengths up to 3.6×10^4 mononucleotides (the size of a single strand in a DNA duplex with mol wt 23.5×10^6) were reliably estimated; in such cases 1.65 ng of terminal phosphorus was detected in a nucleic acid sample of 625 μg. The procedure may well be adapted to DNA with a mass over 10^8 daltons per end-phosphate group. The method involved: (1) enzymatic cleavage of the monoesterified phosphate from the polynucleotide and isolation of the released PO_4^{3-}; (2) simultaneous neutron activation analysis of the total nucleic acid phosphorus and of the fraction hydrolyzed to inorganic PO_4^{3-}. The determinations yielded reference data for calculating the absolute molecular weight of homogeneous polynucleotides and the number-average molecular weight, m_n, of polydispersed nucleic acids. In general, the molecular weight values corresponding to the observed polymerization numbers agreed with these derived by conventional sedimentation measurements. Indirectly, the results demonstrated the experimental validity of a new formula for calculating the m_n of heterogeneous polynucleotides from the molecular weight, m_{max}, of the fraction which was found at the peak of the sedimentation profile during zone ultracentrifugation.

The conventional measurements of the molecular weight, M, of nucleic acids dispersed in aqueous solutions involve: (a) estimate of parameters closely related to their hydrodynamic properties; (b) calculation of M with the aid of empirical equations derived by calibrating the dependent variable *vs.* polynucleotides of known molecular mass. The validity of these data rests ultimately on the accuracy to which the molecular weights of few DNAs and RNAs taken as reference can be determined by an absolute method. In this respect, the most direct approach is undoubtedly sedimentation–diffusion according to the general expression $M = (s^0RT)/D(1 - \bar{v}\rho)$, which is not of practical use since it implies the very difficult evaluations of the diffusion coefficient, D, and of the partial specific volume, $\bar{v}$, of nucleic acids.

The search for reliable experimental correlations between M and the various hydrodynamic attributes of the polynucleotides is almost continual. For instance, Freifelder and coworkers have recently applied several independent techniques to establish definite values of M and s for a set of bacteriophage and virus DNAs which may be regarded as standard molecules (Bancroft and Freifelder, 1970; Dubin *et al.*, 1970; Freifelder, 1970; Lang, 1970).

The degree of polymerization in a nucleic acid can be inferred from the ratio between its total and terminal phosphorus, which is susceptible of being released by phosphomonoesterases and separately assayed as inorganic phosphate. Because of the low resolution, such differential phosphorus measurements have been used only to ascertain the chain length of [^{32}P]oligonucleotides recovered after partial hydrolysis of [^{32}P]DNA (Furlong, 1965) or after neutron activation of nucleic acid digests (Rushizky and Miller, 1967).

The present report deals with the application of the end-phosphorus analysis to highly polymerized nucleotides and outlines a procedure for the distinct counting of total and monoesterified phosphate in DNAs with molecular weights equal to or even exceeding 2.5×10^7 daltons. The method takes advantage of the very sensitive detection of DNA phosphorus by radioactivation (Sabbioni *et al.*, 1971). Absolute data were obtained on either the true molecular weight of homogeneous phage DNA or the number-average molecular weight, m_n, of nucleic acids with a nonuniform chain size.

Materials and Methods

Biochemicals and Reagents. [^{14}C]dCMP was purchased from Schwarz-Mann, $^{32}PO_4^{3-}$ from Amersham Radiochemical Centre, DEAE-cellulose from Whatman, ATP, dATP, and *p*-nitrophenyl phosphate from Sigma, Bio-Gel P-2 and P-6 from Calbiochem, activated charcoal from Merck, and acid aluminum oxide from Woelm (Eschwege, West Germany). Aqueous solutions were prepared with quartz distilled water.

Polynucleotides. Poly(U) was obtained from Schwarz-Mann. Poly(dT) and poly([^{14}C]dT) were prepared with the aid of terminal deoxynucleotidyltransferase as previously described (Bekkering-Kuylaars and Campagnari, 1972). All the polymers were recovered as precipitates in 85% (v/v) aqueous isopropyl alcohol at 2° and then purified on Bio-Gel P-6 columns.

DNA. Homogeneous DNA was isolated from the phage SPP1 of *Bacillus subtilis* as outlined by Riva *et al.* (1968). Polydispersed DNA of nonuniform molecular size was extracted from *B. subtilis* strain 3231 (Polsinelli and Barlati, 1967) by the method of Sarfert and Venner (1965) using a Pronase treatment as a substitution for the 4-aminosalicylate step. [^3H]DNA was obtained from *B. subtilis* grown in the presence of [^3H]thymidine. The DNAs were dissolved in 0.015 M sodium citrate–0.15 M tris-acetate (pH 8.0).

† From the Biology Division (L. C. and F. C.) and the Chemistry Department (E. S. and F. G.), Euratom Joint Research Center, Ispra, Italy, and from the Laboratorio di Genetica Biochimica ed Evoluzionistica del C. N. R., Pavia, Italy (S. S.). *Received August 7, 1972.* This publication is Contribution No. 777 of the Euratom Biology Division.

A series of nucleic acids differing in their average chain length was prepared by sonicating dilute solutions of the bacterial DNA for various time intervals at 2°. The operation was performed at a radio frequence of 20 kHz and a power of 75 W in a Branson Model S-75 apparatus equipped with the 0.5-in. step horn.

The sonicated and nonsonicated samples of DNA were denatured as previously reported (Campagnari *et al.*, 1967) and dialyzed 24 hr at 2° against many changes of Tris-acetate (pH 8.0) whose concentration was progressively raised from 5 mM to 1 M. After dialysis, the concentrations of DNA in solution were estimated by measurements of nucleotide phosphorus (Sabbioni *et al.*, 1971) and adjusted to values of 100–660 μg/ml at an almost inverse proportion with the expected molecular size.

Determination of the Molecular Weight of Nucleic Acids by Sedimentation Analysis. The sedimentation coefficient, *s*, of the nucleic acids was measured by conventional ultracentrifugation and the related molecular weight, *m*, was calculated from the value of s_0 corrected for the solvent at 20°.

About 0.25 μmol of the synthetic polymers were layered over a 3.5–25.0% sucrose gradient in the 1 × 3 in. Strohmaier cells of a Omega II ultracentrifuge (Heraeus-Christ, Osterode, West Germany) supplied with a $\omega^2 dt$ integrator. The fractionating cells were spun for 21 hr at 26,500 rpm and 20°. The median sedimentation distance of the polynucleotides after centrifugation and the corresponding *s* value were determined by the method of Strohmaier (1966).

Band sedimentation of native and denatured DNA (Studier, 1965) was carried out in a Spinco Model E analytical ultracentrifuge equipped with uv optics. Less than 2 μg of DNA were layered in a 12-mm kel-F band-forming centerpiece filled with either 1 M NaCl–1 mM EDTA–10 mM Tris-HCl (pH 8.0) or 0.9 M NaCl–0.1 M NaOH and spun at 39,460 rpm and 20°. After the established velocity was attained, six pictures of the sedimentation chamber were taken at intervals of 2 min. The displacements of the peak of uv-absorbing material were measured in microdensitometric profiles of the photographic plates. $s^0_{20,w}$ was calculated according to Studier (1965).

The following equations were used in the computing of molecular weight: $s^{sucrose}_{20,w} = 0.0117\ m^{0.53}$ for the single-stranded polynucleotides sedimented in a sucrose gradient (Van der Schans *et al.*, 1969); $s^0_{20,w} = 0.0882\ m^{0.846}$ for double-stranded DNA in neutral buffer, $s^0_{20,w} = 0.0105\ m^{0.549}$ and $s^0_{20,w} = 0.0528\ m^{0.4}$ for monohelical DNA in neutral and alkaline buffers, respectively (Studier, 1965).

Phosphomonoesterases. The alkaline phosphatase from *E. coli* was supplied by Worthington and the one from calf intestine by Boehringer. Both preparations were dialyzed for 24 hr at 2° against several changes of 10 mM Tris-acetate (pH 8.0) and then assayed for catalytic activity according to Garen and Levinthal (1960) and for protein concentration assuming $E^{1\,cm}_{1\%} = 7.2$ at 278 nm (Plocke *et al.*, 1962). Specific activities of 18 and 232 units (μmol of *p*-nitrophenyl phosphate hydrolyzed in 1 min by 1 mg of enzyme) were found for the bacterial and the mammalian enzyme, respectively. The phosphatase solutions were diluted to 0.8 mg/ml, frozen, and safely stored at −20° until use; 0.5 ml of these samples were lyophilized and subjected to neutron activation analysis for determination of the P, Zn, and S contents (Bresesti and Neumann, 1964).

Both enzymes were apparently free from DNase contamination since no acid-soluble radioactivity was released from [^{3}H]DNA or poly([^{14}C]dT) incubated for 24 hr at 37° in Tris-acetate (pH 8.0) with 130 μg/ml of one or the other catalytic protein.

Radioactive Labeling of Alkaline Phosphatase. Aliquots of the phosphatase from calf intestine were labeled with either $^{32}PO_4^{3-}$ according to Engstrom (1961a) or ^{65}Zn according to Lazdunsky and Lazdunsky (1969).

Enzymatic Cleavage of Terminal Phosphate from Nucleic Acids. The hydrolysis was carried out at 37° on single-stranded nucleic acids, such as the synthetic polynucleotides and heat-denatured DNA. The reaction mixture (1 ml) contained 1 mmol of Tris-acetate (pH 8.0), 1 unit of phosphatase, and 50–625 μg of the various nucleotides with amounts of monoesterified phosphate ranging from 0.053 to 10.0 nmol. As found by monitoring the produced inorganic PO_4^{3-}, the enzymatic hydrolysis went practically to completion in about 2 hr. However, the incubation was cautiously continued up to 24 hr. Under these conditions, the phosphomonoesterase from calf intestine acted at adequate rates of catalysis also on nucleotides carrying terminal polyphosphate groups. With 4 μmol of either ATP or dATP, the release of PO_4^{3-} was totally accomplished in less than 8 hr, well within the allowed 24-hr incubation.

The hydrolysis was stopped by adding 3 M HNO_3 to the reactant solution up to a final concentration of 0.5 M. The samples were then processed on charcoal columns to separate the formed PO_4^{3-}.

Nucleotide Adsorption on Activated Charcoal. Powdered charcoal was packed in small tubes of polyethylene and extensively washed with 0.5 M HNO_3. Standard charcoal columns of 0.4 × 4.0 cm were used. When loaded with 0.5 M HNO_3 containing either 2.3 μmol of [^{3}H]DNA or 2.0 μmol of [^{14}C]-dCMP, the columns retained more than 99.995% of the radioactivity. Higher concentrations of the acid lowered their adsorption efficiency. In the presence of 0.5 M HNO_3, neither the phosphatase labeled with ^{32}P or ^{65}Zn nor inorganic $^{32}PO_4^{3-}$ were adsorbed on charcoal and more than 99.8% of their radioactivity was recovered in a fraction consisting of the column filtrate plus an acid wash of 1 ml. Only above pH 1.0, measurable amounts of $^{32}PO_4^{3-}$ were retained by the activated charcoal.

Neutron Activation Analysis of the Phosphorus Content in the Samples. Aliquots of the nucleic acid solutions and the eluates from the charcoal columns were dried in silica vials by heating for a few hours at 70°.

A specimen of each nucleotide and two to four PO_4^{3-} fractions produced by separate hydrolyses of its phosphomonoester termini were included in an irradiation capsule of magnesium aluminum alloy (Sabbioni *et al.*, 1971) together with 50 μg of $(NH_4)_2HPO_4$ as a standard.

Sequentially, the samples were subjected to radioactivation of ^{31}P to ^{32}P by thermal neutrons, acid oxydation of all the chemical species of ^{32}P to $^{32}PO_4^{3-}$ absorption of radionuclide contaminants on acid aluminum oxide columns, and measurement of the recovered $^{32}PO_4^{3-}$ by detection of the Cerenkov effect. The details of the procedure were referred in a preceding paper (Sabbioni *et al.*, 1971).

Results

Table I reports the molar content of P, Zn, and S in the phosphomonoesterases from *Escherichia coli* and calf intestine assuming molecular weights of 8 × 10^4 for the former enzyme (Applebury and Coleman, 1969) and 10^5 for the latter one (Engstrom, 1961b). The findings are in agreement with the

TABLE I: Molar Content of P, Zn, and S in Alkaline Phosphatases.[a]

Enzyme Source	P	Zn	S
E. coli	0.23	2.17	
	0.22	2.06	19.3
	0.19	2.12	18.1
	0.21	2.27	
Calf intestine	0.13 (0.16)	1.73 (2.16)	11.6 (14.5)
	0.14 (0.18)	1.78 (2.24)	13.2 (16.5)
	0.13 (0.16)	1.65 (2.06)	
	0.12 (0.15)	1.69 (2.11)	

[a] The data were calculated for the experimental molecular weights of the enzymes from *E. coli* and calf intestine (see text). The values within brackets were obtained when the mammalian preparation was assumed to have the same molecular weight reported for the bacterial phosphatase.

previous data on the alkaline phosphatase from *E. coli* (Plocke *et al.*, 1962; Rothman and Byrne, 1963; Simpson and Vallee, 1968; Applebury and Coleman, 1969; Applebury *et al.*, 1970) and show the similarity between the two proteins.

As reported under Materials and Methods, the phosphorus bound to the phosphomonoesterase hydrolyzing the terminal phosphate of nucleic acids could not be separated on activated charcoal from the released inorganic PO_4^{3-} and actually contributed to raise the ^{32}P blank in the method. Therefore, the amount of phosphatase used in our experiments was lowered as much as possible. In fact, the enzyme from calf intestine was preferred over the bacterial preparation because it contained less phosphorus per unit of catalytic activity.

The total and the terminal phosphorus of various nucleic acids were concomitantly determined by neutron activation analysis. The obtained data represented the relative concentrations of the two ^{32}P fractions and their ratio, actually a ratio between the corresponding counts per minute, allowed to estimate directly the average length of the nucleotide chains in the sample. Practically, the resolution of the method depended upon the quantitative relationship between the absolute amount of the end-phosphate groups contained in a nucleic acid specimen of a definite size and the smallest amount of phosphorus which could be reliably detected. In the present investigation, the minimal aliquot of terminal phosphorus for a reliable measurement was arbitrarily set to 1.5 ng. This value was 20 times larger than the ^{32}P blank contributed by the phosphorylated enzyme plus the overall reagents.

For our experimental production of ^{32}P after 7-days irradiation, the established threshold quantity of phosphorus yielded 300 net cpm as compared with the 24 cpm due to the blank (15 cpm) plus the detector background (9 cpm). When greater amounts of end-phosphate groups were analyzed, the irradiation time could be proportionally shortened. However, care was taken to adjust the size of the nucleic acid samples and their exposure to neutrons for obtaining always more than 300 net cpm in the determination of the terminal phosphorus. Figure 1 records the aliquots of monoesterified phosphate present in 500 μg of monohelical polynucleotides with m_n ranging from 10^3 to 10^7 daltons as a function of the irradiation time calculated to ensure 300 net cpm after a radioactivity decay of 2 days.

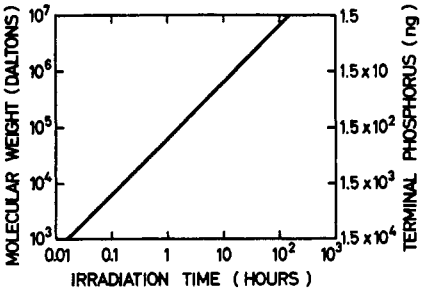

FIGURE 1: Irradiation time yielding 300 net cpm for the terminal phosphorus from 500 μg of various nucleic acids with increasing molecular weights. For details, see text.

Table II lists a series of chain length determinations carried out on various nucleic acid specimens of nonuniform size as well as on a homogeneous phage DNA of known molecular weight. It should be noted that only the amount of SSP1 DNA used for measuring more than 1.5 ng of terminal phosphorus exceeded the 500 μg corresponding to the sample size taken into consideration in Figure 1. The polymerization number of each nucleic acid was multiplied by the mass of its constituent or average nucleotidyl monomer and the resulting $\bar{m}_n$ was compared with the molecular weight estimated from ultracentrifugal measurements.

It should be recalled that the sedimentation method yields a weight-average molecular weight, m_w which, by definition, is equal to m_n for homogeneous polynucleotides and approximates 2 m_n if the sample is polydispersed. In the latter case, moreover, the direct determination of m_w from sedimentation analysis is largely prevented by technical difficulties (see Van der Schans *et al.*, 1969).

As shown in the Appendix of this paper, the m_n of a population of polynucleotides heterogeneous in size may safely be calculated from the formula $m_{max}/m_n = 2 - a$. Here, the notation m_{max} stands for the molecular weight which corresponds to the peak of nucleotide material sedimenting either through a sucrose gradient or in the first phases of analytical zone centrifugation; a represents the exponent of the empirical equation for computing the molecular weight, m, of the nucleic acid under the given conditions (see Materials and Methods).

In all the samples, the difference between repeated measurements of chain length did not exceed a few per cent of the absolute figures. Within such variations, the $\bar{m}_n$'s calculated from these data coincided practically with the m values independently estimated by band sedimentation.

These conclusions applied to the molecular weight determinations carried out either on nucleic acid with molecules of different sizes or on the monodispersed SPP1 DNA. In the former case, the ratios between the m_{max} yielded by ultracentrifugal methods and the m_n calculated from the end-phosphorus analyses approximated the theoretical limit of $2 - a$. A notable exception was the enzymatically synthesized poly(dT), for which the above ratio fell between 1.0 as typical of a set of compounds with the same chain length, and $2 - a$, actually 1.45, as expected for a mixture of heterogeneous molecules. The finding was indicative of the narrow molecular weight distribution in the polymers formed from a given amount of deoxyribonucleoside triphosphates by the terminal transferase (Kato *et al.*, 1967; Chang and Bollum, 1971). The data also demonstrated that ultrasonic vibrations cleaved the phosphodiester bonds of DNA in aqueous solu-

TABLE II: Determination of Molecular Weight in Nucleic Acids. Comparison between End-Phosphate Analyses and Ultracentrifugal Methods.

Nucleic Acid	Chain-Length Measurements: Sample Amt (μg)	No. of Assay	Chain Length $\bar{X}^a \pm \sigma_m{}^b$	$\bar{m}_n{}^c$ ($\times 10^{-3}$)	Molecular Weight Determinations: Method	Cor s^d (S)	$m_{max}{}^e$ or $m_w{}^f$ ($\times 10^{-3}$)	$m_n{}^g = m_{max}/(2-a)^h$ ($\times 10^{-3}$)	m_{max}/m_n or m_w/m_n: Theor Ratio	Exptl Ratio
					Polydispersed Samples					
Poly(dT)	115	3	115 ± 3.5	37.6	Centrifugation in a neutral sucrose gradient	3.46	45.9	31.2	1.47	1.22
Poly(U)	120	2	210 ± 2.6	69.1		5.26	101.0	68.7	1.47	1.46
DNA I[i]	110	4	291 ± 6.5	96.3	Band sedimentation in neutral buffer (heated DNA)	6.91	136.0	93.8	1.45	1.41
DNA II[i]	165	4	722 ± 17.3	239.0		11.00	317.2	218.8	1.45	1.33
DNA III[i]	220	4	1,619 ± 51.2	535.9	Band sedimentation in alkaline buffer	12.22	815.0	509.4	1.60	1.52
DNA IV[i]	330	4	3,356 ± 49.7	1,108.4		16.42	1,705.0	1,065.6	1.60	1.54
DNA V[i]	495	4	32,900 ± 823	10,889.9		39.70	13,502.0	9,688.8	1.60	1.42
					Homogeneous Sample					
SPP1 DNA	625	3	35,567 ± 641	11,772.7	Band sedimentation in alkaline buffer	36.32 ± 0.75[j]	12,430.0		1.0	1.06
					Band sedimentation in neutral buffer (native DNA)	31.94 ± 0.54[j]	24,860.0 (24,970.0[k])		2.0	2.12 (2.12)

[a] $\bar{X} = \Sigma X_i/n$ = mean chain length value. [b] σ_m = standard error = $\sigma/(n)^{1/2}$, where σ is equal to $[\Sigma(\bar{X} - X_i)^2/(n-1)]^{1/2}$. [c] $\bar{m}_n$ = number-average molecular weight calculated from $\bar{X}$. [d] s = sedimentation coefficient. [e] m_{max} = molecular weight of the polydispersed nucleotide material at the peak of the sedimentation profile. [f] m_w = molecular weight (weight average) of homogeneous polynucleotides = $\Sigma n_i m_i{}^2/\Sigma n_i m_i$. [g] m_n = number-average molecular weight = $\Sigma n_i m_i/\Sigma n_i$. [h] a = exponent of empirical relations $s = km^a$. [i] DNAs from *B. subtilis* sonicated 12 min (I), 4 min (II), 50 sec (III), 20 sec (IV), and not sonicated (V). [j] Standard error of $s^0_{20,w}$ (three determinations). [k] Calculated according to the Freifelder (1970) equation. For details, see Materials and Methods and Results Sections.

tions, thus producing the same number of polynucleotide fragments and phosphoryl end groups.

During ultracentrifugation, the SPP1 DNA sedimented as a single species, being always confined in a narrow band without notable dispersion. The m_w values were 24.90×10^6 in neutral and 12.43×10^6 in alkaline buffers (Table II) and agreed with the data originally obtained by the same methods (Riva *et al.*, 1968). The results showed the absence of monohelical breaks in the preparation. Assuming the presence of one monoesterified phosphate is each strand of the DNA duplex, the $\bar{m}_n$ value derived from the end-phosphorus determination should have been equal to either the m_w in alkali or 0.5 m_w at pH 8.0. Indeed, the found $\bar{m}_n$ approximated the m_w values of the monohelical nucleic acid, being only 5% lower. As seen in Table II, the difference between such data remained practically unchanged when the m_w of the double-stranded DNA was computed according to the new equation $s^0_{20,w} = 2.8 + 0.00834\ m^{0.479}$ proposed by Freifelder (1970) on the basis of more correct experimental values. The overall findings on the SPP1 DNA proved the existence of terminal phosphoryl groups in the helices of the molecule. It should also be noted that the precision and the accuracy of the end-phosphorus measurements were not below those of the $s^0_{20,w}$ estimates by analytical ultracentrifugation.

Discussion

The described determinations of chain length in nucleic acids may yield reference data for calibrating their hydrodynamic parameters from which molecular weight values can be derived.

By adjusting concurrently a few experimental conditions, the resolution of the method may easily be enhanced by one order of magnitude. For instance, the assigned limit to the detection of terminal phosphorus can safely be lowered by a factor of 2. Moreover, the time of neutron activation can be double up to the 14 days of the ^{32}P half-life with a proportional increase of the actual counts per minute, while the amounts of nucleotide analyzed for the phosphate end groups can be conveniently augmented. Then, it should be possible to determine the chain length of a DNA duplex with mol wt 25×10^7 and two monoesterified phosphates per molecule from a sample of 3 mg. After 2-weeks irradiation and 2-days radioactive decay, the 0.75 ng of terminal phosphorus would result in 300 net cpm over the 40 cpm due to the ^{32}P blank plus the background of the scintillation spectrometer.

The use in our experiments of the very active alkaline phosphatase from calf intestine was essential because it allowed to minimize the aliquot of enzyme protein needed to hydrolyze the monoesterified phosphate. Consequently, the contamination of the end-phosphorus fraction by the phosphoryl groups present in the common phosphomonoesterases at an almost constant ratio (see Results, Table I) was greatly reduced. An amount of the *E. coli* enzyme with comparable phosphatasic activity would have enhanced the ^{32}P blank more than ten times with a proportional loss in the sensitivity of the terminal phosphorus measurements.

The reported determinations of chain length were performed on originally single-stranded or thermally denatured nucleic acids carrying a phosphoryl terminus. However, the conditions of the assays could easily be adapted to DNA duplexes with monoesterified phosphates at the ends and/or in the monohelical breaks of the molecule. Indeed, alkaline phosphatase was shown to release specifically external and internal phosphomonester groups from native DNA depending upon the incubation temperature and, possibly, the ionic strength of the medium (Weiss *et al.*, 1968).

The outlined procedure may even be applied to detect the polymerization number of nucleic acids which have diphosphate or triphosphate termini as the 5S RNA from HeLa cells (Halten *et al.*, 1969) and wheat germs (Soave *et al.*, 1970). In fact, the alkaline phosphatase from calf intestine was able to split the polyphosphate end groups from ribonucleotides (Fernley and Walker, 1967) as well as from deoxyribonucleotides (see Materials and Methods).

Finally, the equivalence in the polydispersed nucleic acids of the m_n values obtained from the chemical and physical measurements might be viewed in reverse as an experimental proof of the new formula $m_n = m_{max}/(2 - a)$. The proposed equation, mathematically derived and considered in the Appendix, was used to calculate m_n from the m_{max} yielded by the sedimentation analyses.

Appendix: Distribution of Molecular Weight and Sedimentation Coefficient in Zonal Centrifugation of DNA Samples with Molecules of Random Chain Length[1]

Govert Paulus van der Schans and Francesco Campagnari[1]

It was shown by Van der Schans *et al.* (1969) that random breakage of DNA molecules of molecular weight, M gives a distribution $P(m)dm$ of molecular weights (m) (unbroken

$$P(m)dm = (p\,dm/M)\left\{2 + \left(1 - \frac{m}{M}\right)p\right\}\exp(-pm/M) \quad (1)$$

molecules not included), where p is the average number of breaks per molecule.

Analytical band centrifugation of monodisperse DNA molecules broken randomly, results in a sedimentation profile. The distribution function $E(x)$ of the extinction E as a function of distance x from the axis of rotation can be written as

$$CE(x) = mP(x) = mP(m)\frac{dm}{dx} = mP(m)\frac{dm}{ds}\frac{ds}{dx} \quad (2)$$

where C is a normalization factor and s the sedimentation coefficient, while dm/ds can be derived from the relation $s = km^a$.

The sedimentation coefficient s is related to the distance x from the axis of rotation by the well-known relation

$$s = \left(\frac{dx/dt}{\omega^2 x}\right)b$$

where b is a constant and ω is the angular velocity. From the integrated form

$$x = x_m \exp\left(\frac{\omega^2 st}{b}\right)$$

[1] From the Medical Biological Laboratory TNO, Rijswjk Z. H., The Netherlands (G. P. V.), and from the Biology Division, Euratom Joint Research Center, Ispra, Italy (F. C.). This publication is Contribution No. 776 of the Euratom Biology Division. The authors wish to thank Dr. L. Farese for useful discussion on the theoretical aspect of this study.

where x_m is the distance between meniscus and axis of rotation, ds/dx can be derived. Using the relations for dm/ds and ds/dx, and eq 2, it follows from eq 1 that

$$CE(x) = \frac{m^{2-a}}{M}p\left\{2 + \left(1 - \frac{m}{M}\right)p\right\} \times \exp(-pm/M)\exp\left(-\frac{k\omega^2 t m^a}{b}\right)\frac{b}{kax_m\omega^2 t} \quad (3)$$

If p is known, $E(x)$ can be computed as a function of x at the time t. As it may be proved by elementary calculus, $E(x)$ reaches a maximum at a certain value of p if

$$(4 + 2p - 2a - ap)m^{-a} + (a - 5 - p + pm/M)pm^{1-a}/M - \frac{ak\omega^2 t}{b}(2 + p - pm/M) = 0 \quad (4)$$

The DNA molecular weight, $m = m_{max}$, measured at the maximum of $E(x)$, is clearly a function t. However, this is not a serious problem. When one takes into consideration the empirical relation $s^0_{20,w} = 0.053m^{0.4}$ for the sedimentation of DNA in alkali according to Studier (1965) and assumes that the polynucleotide fragments were produced by introducing ten random breaks per molecule of molecular weight M, one may substitute 0.053, 0.4, and 10 for k, a, and p, respectively. Under these conditions with short centrifugation times up to 600 sec and with $\omega^2/b = 1.707 \times 10^7/1.16 \times 10^{13} = 1.47 \times 10^{-7}$, the m_{max} calculated from eq 4 is nearly independent of t, *i.e.*, m_{max} can be approximated by substituting $t = 0$ in eq 4. This gives

$$\frac{m_{max}}{M} = \frac{p + 5 - a}{2p} - \frac{1}{2p}\{p^2 + (2a + 2)p + (9 - 2a + a^2)\}^{1/2} \quad (5)$$

This equation is exactly the same as that derived from the distribution function $E(s)$ for sucrose gradients (Van der Schans *et al.*, 1969). The relevant problem is that of the meaning and of the real nature of m_{max}. This value is distinct from the number-average molecular weight m_n which is defined as $\Sigma n_i m_i/\Sigma n_i$ and which is here equal to $M/(p + 1)$. If eq 5 is divided by m_n/M, we have

$$\frac{m_{max}}{m_n} = (p + 1)\left[\frac{p + 5 - a}{2p} - \frac{1}{2p}\{p^2 + (2a + 2)p + (9 - 2a + a^2)\}^{1/2}\right] \quad (6)$$

For large values of p and practically at $p > 10$, m_{max}/m_n approximate $2 - s$. The sedimentation coefficient s_{max} of the nucleic acid fraction corresponding to the peak of $E(x)$ allows to estimate m_{max} from a simple equation of the general kind $s = km^a$ and thus to obtain directly m_n from $m_{max}/(2 - a)$.

The reported equations were derived for the case of nucleic acid fragments produced by random degradation of an original population of molecules with equal length. As shown by Charlesby (1960), a most probable distribution of molecular size is obtained when a large number of breaks are delivered over a set of linear polymers with identical chain length. However, a size distribution of the most probable type is reasonably maintained also in the course of random breakage of heterogeneous nucleic acid samples whose parental molecules initially obeyed a most probable distribution of molecular dimensions. Therefore the relation $m_{max}/m_n = 2 - a$ can be used in band sedimentation analyses of polymucleotides to calculate the number-average molecular weight of nonuniform molecular populations with random chain length.

The extinction profile of the DNA which sediments during zone centrifugation records the nucleotide concentration in the serial cross sections of the rotating cell. The $E(x)$ pattern coincides with the real mass distribution only if the transverse area of the sedimentation chamber is uniform and no loss of material on the cell wall occurs.

When centrifugation is carried out in sector-shaped tubes, corrections for radial dilution have to be included. These corrections have not been taken into account while deriving the above equations since the amendment is often of little relevance or is readily combined with the numerical computation of the $E(x)$ maximum. Actually, the $E(x)$ distortion due to radial dilution may well be neglected when the bandwidth and the sedimentation of the DNA are small in comparison with the distance to the axis of rotation as in the band ultracentrifugation analyses. Otherwise, the space interval covered by the sedimented polynucleotides in the sucrose gradients is quite large. However, the radial dilution effect is easily corrected for by plotting the amount of the nucleic acid and not its concentration versus the sedimentation distance as suggested by Strohmaier (1966) and practised in the experiments reported under Results (see also Methods).

References

Applebury, M. L., and Coleman, J. E. (1969), *J. Biol. Chem. 244*, 308.

Applebury, M. L., Johnson, B. P., and Coleman, J. E. (1970), *J. Biol. Chem. 245*, 4968.

Bancroft, F. C., and Freifelder, D. (1970), *J. Mol. Biol. 54*, 537.

Bekkering-Kuylars, S. A. M., and Campagnari, F. (1972), *Biochim. Biophys. Acta 272*, 526.

Bresesti, M., and Neumann, H. (1964), Report Euratom, 1601.

Campagnari, F., Bertazzoni, U., and Clerici, L. (1967), *J. Biol. Chem. 242*, 2168.

Chang, L. M. S., and Bollum, F. J. (1971), *Biochemistry 10*, 536.

Charlesby, A. (1960), Atomic Radiation and Polymers, New York, N. Y., Pergamon Press, Chapter 10.

Dubin, S. B., Benedek, G. B., Bancroft, F. C., and Freifelder, D. (1970), *J. Mol. Biol. 54*, 547.

Engstrom, L. (1961a), *Biochim. Biophys. Acta 52*, 49.

Engstrom, L. (1961b), *Biochim. Biophys. Acta 52*, 36.

Fernley, N. H., and Walker, P. G. (1967), *Biochem. J. 104*, 1011.

Freifelder, D. (1970), *J. Mol. Biol. 54*, 567.

Furlong, N. B. (1965), *Anal. Biochem. 12*, 349.

Garen, A., and Levinthal, C. (1960), *Biochim. Biophys. Acta 38*, 470.

Halten, L. E., Amaldi, E., and Attardi, G. (1969), *Biochemistry 8*, 4989.

Kato, R., Goncalves, J. M., Houts, G. E., and Bollum, F. J. (1967), *J. Biol. Chem. 242*, 2780.

Lang, D. (1970), *J. Mol. Biol. 54*, 557.
Lazdunski, C., and Lazdunsky, M. (1969), *Eur. J. Biochem. 7*, 294.
Plocke, D. J., Levinthal, C., and Vallee, B. L. (1962), *Biochemistry 1*, 373.
Polsinelli, M., and Barlati, S. (1967), *J. Gen. Microbiol. 49*, 267.
Riva, S., Polsinelli, M., and Falaschi, A. (1968), *J. Mol. Biol. 35*, 347.
Rothman, F., and Byrne, R. (1963), *J. Mol. Biol. 6*, 330.
Rushinzky, G. W., and Miller, W. W. (1967), *Anal. Biochem. 20*, 181.
Sabbioni, E., Clerici, L., Campagnari, F., Girardi, F., and Bartolini, P. (1971), *Biochemistry 10*, 4185.
Sarfert, E., and Venner, H. (1965), *Hoppe-Seyler's Z. Physiol. Chem. 340*, 157.
Simpson, R. T., and Vallee, B. L. (1968), *Biochemistry 7*, 4343.
Soave, C., Galante, E., Torti, G. (1970), *Bull. Soc. Chem. Biol. 52*, 857.
Strohmaier, K. (966), *Anal. Biochem. 15*, 109.
Studier, F. W. (1965), *J. Mol. Biol. 11*, 373.
Van der Schans, G. P., Aten, G. B. T., and Blok, J. (1969), *Anal. Biochem. 32*, 14.
Weiss, B., Live, T. R., Richardson, C. C. (1968), *J. Biol. Chem. 243*, 4530.

Properties of Oligodeoxynucleotides That Determine Priming Activity with *Escherichia coli* Deoxyribonucleic Acid Polymerase I†

Mehran Goulian,* Susan H. Goulian, Ellen E. Codd, and Aija Z. Blumenfield

ABSTRACT: The requirements for priming of DNA synthesis by oligonucleotides with *Escherichia coli* DNA polymerase I and single-stranded templates have been examined with respect to sequence and chain length. Sequences containing only pyrimidines have very little activity, whereas purine sequences are highly active. The highest activity in an oligonucleotide of defined sequence was found with $d(pApG)_4$. Maximum priming activity at 20° occurs with chain lengths of 8–12, but activity can be detected even with trimers and tetramers, suggesting that the enzyme plays a role in stabilizing the association between template and primer. Oligo(dG) had only slight activity, whereas there was no activity for the homo-oligonucleotides of dA, dT, or dC. Some requirements for activity are probably based on general features of oligonucleotide–polynucleotide interactions, but it is not clear that all of the requirements have this explanation.

Studies (*in vitro*) with *Escherichia coli* DNA polymerase I have established that the purified enzyme is unable to initiate new polynucleotide chains (Goulian, 1968a). The enzyme appears to commence synthesis only by extension of existing chains at the 3′-OH termini. Oligodeoxynucleotides can provide this priming function with single-stranded templates (Goulian 1968a), an observation that had been made earlier for animal cell DNA polymerases (Keir, 1962; Smellie, 1963; Furlong, 1966; Bollum, 1967). It is postulated that the oligonucleotides associate with template in a short region of sequence homology, and nonpaired residues at the 3′-OH terminus are trimmed back by the 3′→5′ exonuclease associated with DNA polymerase I, providing thereby a suitable 3′-OH priming site for the enzyme. The oligonucleotides are incorporated into the product in covalent form although as synthesis proceeds they are largely removed, presumably by the 5′→3′ nuclease associated with polymerase I (Goulian, 1968b).

Preliminary studies on the properties of oligonucleotide primers for *E. coli* DNA polymerase I suggested little or no specificity; fragments active in priming could be produced by digestion of DNA from several different sources with pancreatic DNase. It also appeared that fragments could be quite short and still retain primer activity although this was not firmly established.

In the present study the requirements for primer activity are defined in greater detail, with specific attention to the influence of chain length, base sequence, nucleotide sugar, enzyme, and incubation conditions.

Materials and Methods

Single Strands of DNA

Escherichia coli. *E. coli* 15T⁻ was grown at 37° with aeration in M-9 medium (Adams, 1959) containing, per milliliter, [³H]thymidine, 0.5 μg (1 μCi/μg). The cells were lysed by the lysozyme–Brij procedure of Cozzarelli *et al.* (1968). After centrifugation (60,000*g*, 15 min) and removal of the supernatant containing the minicircular plasmid DNA, the pellet DNA was resuspended in 50 mM EDTA–1% Sarkosyl and incubated with Pronase (1 mg/ml) for 18 hr, 37°. The mixture was centrifuged in 12-ml alkaline sucrose gradients in the

† From the Department of Medicine, University of California—San Diego, La Jolla, California 92037, and the Departments of Medicine and Biochemistry, University of Chicago, and Argonne Cancer Research Hospital, Chicago, Illinois (operated by the University of Chicago for the U. S. Atomic Energy Commission). *Received September 5, 1972.* This research was supported in part by the U. S. Public Health Service (Grant No. 5-RO1-CA11705), the American Cancer Society (Grant No. P-567), and the U. S. Atomic Energy Commission (Argonne Cancer Research Hospital, University of Chicago).