Nucleotide Clusters in Deoxyribonucleic Acids. Comparison of the Sequences of the Large Pyrimidine Oligonucleotides of Bacteriophages S13 and φX174 Deoxyribonucleic Acids†

Barbara Harbers, Allen D. Delaney,‡ Klaus Harbers, and John H. Spencer*

ABSTRACT: The large pyrimidine oligonucleotides from the DNAs of the two related bacteriophages φX174 and S13 have been sequenced. The largest pyrimidine oligonucleotide present is unique to S13 DNA and is the undecanucleotide C_5T_6, sequence C-T-T-C-C-T-C-T-T-C-T. Considerable sequence homology has been found between the pyrimidine oligonucleotides of the two phage DNAs. Out of 14 oligonucleotide sequences from S13 DNA (120 bases) at least ten are identical with sequences of oligonucleotides from φX174 DNA (92 bases) and two are closely related (17 bases), the only difference being a single thymine to cytosine transition in each sequence (a total of 107 identical bases). The pyrimidine oligonucleotides of each phage DNA show extensive internal sequence homology among each other with up to eight bases identical in sequence in pairs of different oligonucleotides. Another interesting observation is the occurrence of symmetrical sequences (true palindromes) which read the same forwards as backwards. The longest symmetrical sequence is the nonanucleotide C_4T_5 sequence, C-T-C-T-T-T-C-T-C, present in both S13 and φX174 DNAs. The extensive sequence homology observed between the pyrimidine oligonucleotides of S13 and φX174 supports the close relationship of the two phages and provides further evidence that they were derived from recent common ancestors.

The small icosahedral bacteriophages S13 and φX174 both contain single-stranded circular DNAs (Tessman, 1959; Sinsheimer, 1959; Fiers and Sinsheimer, 1962; Spencer et al., 1972) which are almost identical in size and physical characteristics. The phages are serologically related (Zahler, 1958), the number of genes the same and the size of the cistrons similar (Benbow et al., 1971; Baker and Tessman, 1967; Tessman, 1965) as is the pattern of phage-coded proteins separated on sodium dodecyl sulfate polyacrylamide gels (Jeng et al., 1970; Godson, 1973). In mixed infections the two phages with the exception of one gene genetically complement (Jeng et al., 1970) and recombine (Tessman and Schleser, 1963). Recently Godson (1973) described a heteroduplex analysis of the two DNAs and reported that despite these similarities only 4.7 ± 1.9% of φX174 DNA is highly homologous with S13 DNA.

The large number of mutants available of both φX174 and S13 phages and the recent description of base sequences from some regions of the φX174 genome (Galibert et al., 1974; Barrell et al., 1975) make these two phages a very useful system for studying the limits of base sequence mismatch tolerated in genetic recombination and DNA heteroduplex experiments. In this report the base sequences of the large pyrimidine oligonucleotides from S13 and φX174 DNA are presented. It is shown that considerable sequence homology exists between the pyrimidine oligonucleotides of the two phages. All the oligonucleotides sequenced which have the same base composition and are present in both phages have identical sequences.

Materials and Methods

Chemicals and Enzymes. All chemicals used were of reagent grade. DEAE-Sephadex A-25 was purchased from Pharmacia (Canada) Ltd. DEAE-cellulose and cellulose for thin-layer chromatography were products of Macherey, Nagel and Co. Cellulose acetate strips were purchased from Schleicher and Schuell Inc., Keene, N.H., and cellogel strips manufactured by Chemetron, Milan, Italy, from Mandel Scientific Co. Enzymes spleen phosphodiesterase, snake venom phosphodiesterase, and bacterial alkaline phosphatase (electrophoretically pure) were obtained from Worthington Biochemical Corp. Alkaline phosphatase was heat treated to remove phosphodiesterases (Garen and Levinthal, 1960); snake venom phosphodiesterase was pretreated at 37°C for 3 hr at pH 3.6 to remove 5′-nucleotidase activity (Sulkowski and Laskowski, 1971). Hexokinase was purchased from Boehringer Mannheim, Germany, and myokinase and lysozyme were from Sigma Chemical Co., St. Louis, Mo. $H_3{}^{32}PO_4$ was obtained from New England Nuclear (Canada). [γ-^{32}P]ATP was prepared as described by Schendel and Wells (1973).

Polynucleotide kinase was prepared from *Escherichia coli* B infected with bacteriophage T4 n82 (Warner and Lewis, 1966) by the method of Richardson (1965) with the following modifications. All buffers in the preparation, until the enzyme was loaded to the first DEAE-cellulose column, contained 1 m*M* ATP. Instead of a stepwise elution from the DEAE-cellulose column a 2-l. linear gradient of 10 m*M* potassium phosphate buffer (pH 7.5)–10 m*M* mercaptoethanol to 50 m*M* potassium phosphate buffer (pH 7.5)–10 m*M* mercaptoethanol was used. The phosphocellulose column was eluted with a 1-l. linear gradient of 0–0.5 *M* KCl, in 50 m*M* potassium phosphate buffer (pH 7.5)–10 m*M* mercaptoethanol. Final purification of the enzyme was obtained by chromatography on a second DEAE-cellulose column as described by Wu and Kaiser (1967). This column

† From the Department of Biochemistry, McGill University, Montreal, Quebec, Canada H3G 1Y6. *Received July 8, 1975.* This is paper XI of a series. Paper X is Harbers et al. (1974). This work was supported by Grant MT 1453 from the Medical Research Council of Canada.

‡ Holder of a studentship from the Medical Research Council of Canada, and Conseil de la Recherche en Santé du Québec.

separated the polynucleotide kinase into two fractions, the first eluting with 10 m*M* potassium phosphate buffer (pH 7.5) and the second eluting with 50 m*M* potassium phosphate buffer (pH 7.5). The first fraction was free of phosphomonoesterase, exonuclease, endonuclease, and deaminase activities and was used in all experiments described in this paper. The assay for polynucleotide kinase used a mixture of terminally dephosphorylated pyrimidine oligonucleotides longer than hexanucleotides as substrate. These were prepared by depurination of the calf thymus DNA with formic acid–diphenylamine (Burton, 1967) followed by dephosphorylation with alkaline phosphatase. The hydrolysate was applied to a DEAE-cellulose column and oligonucleotides containing less than seven bases were eluted with 0.18 *M* NaCl and discarded. The column was then washed with 100 ml of 50 m*M* triethylammonium bicarbonate (TEAB) (pH 8.0) to remove the NaCl and finally the oligonucleotides longer than seven bases eluted with 1 *M* TEAB. The buffer was evaporated in vacuo; the oligonucleotides were dissolved in water and stored at −20°C. The assay was performed in a volume of 0.1 ml of 10 m*M* Tris-HCl buffer (pH 8.1) containing 10 m*M* $MgCl_2$, 20 m*M* mercaptoethanol, 2 nmol of oligonucleotides, and 0.5 nmol of [γ-^{32}P]ATP. Incubation was for 30 min at 37°C. DEAE-cellulose suspended in water was added to the incubation mixture and the suspension filtered through a glass fiber filter. Unreacted [γ-^{32}P]ATP was eluted from the filter by washing with 0.18 *M* NaCl and the radioactivity of the ^{32}P-labeled oligonucleotides on the filter determined by measurement of the Cerenkov radiation in a Beckman LS250 scintillation spectrometer. One unit of enzyme activity is the amount of enzyme which catalyzes the transfer of 1 nmol of phosphate per minute under these assay conditions.

Purification of Unlabeled and ^{32}P-Labeled Bacteriophage S13 Wild Type DNA. This was performed according to Spencer and Boshkov (1973). From a 1-l. culture 2 mg of pure DNA was obtained with a specific activity of 25000 dpm/μg.

Purification of Unlabeled and ^{32}P-Labeled Bacteriophage ϕX174am3 DNA. E. coli C was grown in 500 ml of mT3XD medium (Denhardt, 1969) for preparation of unlabeled DNA and TPG medium (Denhardt et al., 1967) for ^{32}P-labeled DNA. When the cell concentration reached 2 × 10^8/ml the culture was inoculated with bacteriophage ϕX174*am*3 (kindly supplied by Dr. D. T. Denhardt) at a multiplicity of five phage per cell. For ^{32}P-labeled preparations 40 μCi/ml of $H_3{}^{32}PO_4$ was added at the same time. Growth was continued for 2 hr at 37°C, and then the cells were collected by centrifugation and resuspended in 40 ml of 50 m*M* Tris-HCl–10 m*M* EDTA (pH 8.1). The phage was released by incubation with lysozyme (0.1 mg/ml) for 30 min at 37°C. The suspension was then diluted with the Tris-HCl buffer (pH 8.1) to 150 ml, made 0.5 *M* in NaCl, and stirred vigorously at 4°C for 30 min. Cell debris was removed by centrifugation and the phage particles were precipitated by addition of poly(ethylene glycol) (Carbowax 6000, Union Carbide) to the supernatant to a final concentration of 10% (w/v). The solution was stirred overnight at 4°C (Yamamoto et al., 1970). The precipitate was collected by centrifugation at 16000*g* for 20 min (10,000 rpm, Sorvall GSA rotor) and resuspended in 10 ml of 50 m*M* sodium tetraborate (pH 9.2). The suspension was kept at 4°C for at least 1 hr and then clarified by centrifugation for 10 min at 12000*g* (10000 rpm, Sorvall SS34 rotor). Final purification of the phage was accomplished by CsCl density-gradient centrifugation with a starting density of 1.42 g/cm^3. The phage band was collected dropwise from the bottom of the centrifuge tube and diluted and the phage particles were pelleted by centrifugation for 1 hr at 58000 rpm in a Beckman 60 Ti rotor. The phage DNA was released by hot phenol extraction (Sinsheimer, 1966) and precipitated with 2-propanol. The yield was 1 mg of DNA. ^{32}P-labeled DNA had a specific activity of 100000 dpm/μg.

Depurination of DNA to Pyrimidine Oligonucleotides. The DNA from bacteriophages S13 and ϕX174 was depurinated by treatment with 2% diphenylamine in 67% formic acid at 30°C for 18 hr (Burton, 1967). The hydrolysate was extracted three times with ether, and dried in a desiccator under vacuum.

5′-Terminal Labeling of Pyrimidine Oligonucleotides. Depurinated DNA (100 μg) was incubated with 20 μg of alkaline phosphatase in 500 μl of 50 m*M* Tris-HCl–10 m*M* $MgCl_2$ (pH 8.9) at 55°C for 1 hr. The enzyme was removed by phenol extraction, the phenol removed by ether extraction, and the aqueous extract dried and redissolved in 500 μl of a solution containing 10 μmol of mercaptoethanol and 0.03 unit of polynucleotide kinase. [γ-^{32}P]ATP was added in a two- to tenfold excess over the amount of oligonucleotide substrate. Incubation was for 6–10 hr at room temperature. Excess [γ-^{32}P]ATP was converted to glucose 6-phosphate by incubation with 100 μmol of glucose, 100 μg of hexokinase, and 100 μg of myokinase (Hänggi et al., 1970). The reaction mixture was applied to a 1 × 5 cm DEAE-Sephadex A-25 column, the glucose 6-phosphate eluted with 0.2 *M* TEAB (pH 8.0) and the 5′-labeled oligonucleotides eluted with 1 *M* TEAB.

In some experiments the excess ATP was not converted to glucose 6-phosphate but eluted from the DEAE-Sephadex A-25 column with 0.4 *M* TEAB. Under these conditions some of the shorter oligonucleotides are eluted with the ATP and are missing from the subsequent chromatographic separations of the 5′-^{32}P-labeled oligonucleotide material eluted by 1 *M* TEAB.

Fractionation of Pyrimidine Oligonucleotides. The pyrimidine oligonucleotides released by depurination of S13 and ϕX174 DNA were fractionated by the ionophoresis-homochromatography thin-layer system of Brownlee and Sanger (1969) as modified for pyrimidine oligonucleotides by Ling (1972a). The depurination products were first applied to a 3 × 50 cm cellulose acetate or cellogel strip (Trim and Dickerson, 1974), and separated by electrophoresis at 4.5 kV for 20–30 min in a buffer containing 4.5% acetic acid, 0.5% formic acid, 7 *M* urea, and 5 m*M* EDTA (pH 3.5). The separated oligonucleotides were then transferred onto a thin-layer plate coated with a mixture of DEAE-cellulose and cellulose in a ratio of 1:7.5. Chromatography was carried out with a 2% solution of partially hydrolyzed yeast RNA at 60°C (Jay et al., 1974). The separated oligonucleotides were visualized by radioautography and when necessary the spots scraped from the plate and the oligonucleotides eluted with 1 *M* TEAB (pH 8.0).

In some experiments the depurination products were separated by column chromatography as described by Černý et al. (1968, 1969). No carrier DNA was added. For better separation of small amounts of oligonucleotides DEAE-Sephadex A-25 was used instead of DEAE-cellulose. Fractionation according to chain length was performed either directly after depurination using a linear gradient of 0–0.4 *M* NaCl in 7 *M* urea–0.1 *M* sodium acetate (pH 5.5) or

after dephosphorylation of the depurination products using a linear gradient of 0–1.3 *M* TEAB (pH 8.0). Fractionation according to base composition was by standard methods (Černý et al., 1969) but using DEAE-Sephadex A-25. Assignment of the base composition of the oligonucleotides was by means of a grid, in which the chain length and base composition of each component is plotted vs. the molarity of ammonium formate which eluted the component from the DEAE-Sephadex A-25 column at pH 3.0 (Harbers and Spencer, 1974).

Partial digestions of oligonucleotides with spleen and snake venom phosphodiesterase were performed as described by Ling (1972a). Carrier RNA (50–100 μg) was added prior to digestion of oligonucleotides that had been isolated by column chromatography. No differences were observed when commercially available spleen phosphodiesterase or the enzyme isolated according to Bernardi and Bernardi (1968) was used. Digestion products were separated by ionophoresis–homochromatography as described above.

Results

Figure 1a and c show radioautograms of uniformly ^{32}P-labeled pyrimidine oligonucleotides present in formic acid–diphenylamine hydrolysates of S13 and ϕX174 DNAs, respectively, separated by ionophoresis–homochromatography. Studies by Ling (1972a) have shown that the oligonucleotides separate according to base composition in the electrophoretic step, and according to chain length in the chromatography step, thus the relative position of the separated oligonucleotides to each other is composition dependent. This allows the accurate prediction of the base composition of a pyrimidine oligonucleotide by its position on the chromatogram. Base composition grids derived from the pyrimidine oligonucleotide patterns 1a and 1c are shown in Figure 1b and d, respectively. The oligonucleotides present in Figure 1a and c are represented by black dots on the respective grids 1b and d. Differences in occurrence and distribution of the longer pyrimidine oligonucleotides in the two DNAs are immediately apparent. The longest pyrimidine oligonucleotide in S13 DNA is the undecanucleotide C_5T_6 whereas in ϕX174 DNA the longest are the decanucleotides C_2T_8 and C_7T_3. The oligonucleotides C_5T_6, C_6T_4, C_5T_2, and T_7 are present only in S13 DNA, whereas ϕX174 DNA contains C_7T_3 which is absent from S13 DNA.

Pyrimidine oligonucleotides isolated from formic acid–diphenylamine hydrolysates of DNA have 3′ and 5′ terminal phosphates. In some experiments pyrimidine oligonucleotides were released from unlabeled DNA by depurination, then the terminal phosphates were removed, and a radioactive ^{32}P phosphate group was incorporated at the 5′-terminus using polynucleotide kinase. Figure 1e shows the radioautogram and Figure 1f the corresponding grid pattern of 5′-^{32}P terminally labeled oligonucleotides from ϕX174 DNA separated by ionophoresis–homochromatography. Some of the short oligonucleotides are not present because they were removed together with the excess radioactive ATP by column chromatography (see Materials and Methods) prior to electrophoresis. The different grid pattern of Figure 1f compared to that of Figure 1d provided a cross check of the base composition assignments of the longer oligonucleotides. This was particularly important because the assignments of some of the nona- and decanucleotides did not correspond to those previously reported by Ling (1972b).

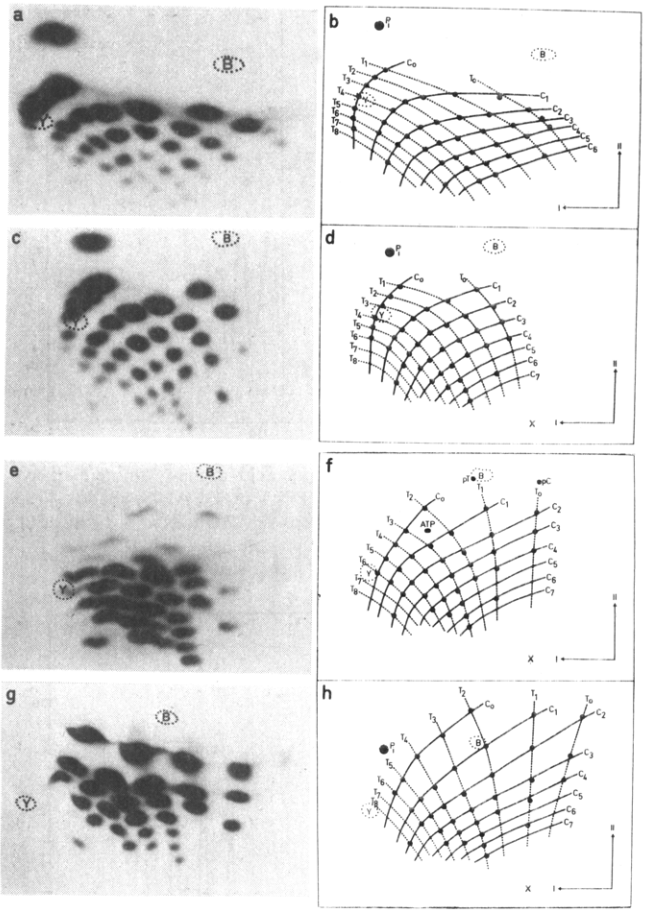

FIGURE 1: Two-dimensional fractionation of the pyrimidine oligonucleotides of bacteriophage S13 wild type and ϕX174*am*3 DNAs. Dimension I, ionophoresis at pH 3.5. Dimension II, homochromatography on DEAE-cellulose thin-layer plates at 60°C, eluent 2% partially hydrolyzed yeast RNA containing 7 *M* urea. The separated oligonucleotides were visualized by radioautography. Y, position of yellow dye marker (orange G); B, position of blue dye marker (xylene cyanol FF); P_i, inorganic phosphate; X, origin. (a) Pyrimidine oligonucleotides released from uniformly ^{32}P-labeled S13 DNA by formic acid–diphenylamine hydrolysis. (b) Grid pattern of 3′,5′ terminally phosphorylated pyrimidine oligonucleotides. The pattern is derived from the oligonucleotide separation of Figure 1a. The position of each oligonucleotide is represented by a dot. Solid lines join oligonucleotides of similar cytosine content, broken lines those of similar thymine content. Origin is outside the area of the chromatogram shown here. (c) Pyrimidine oligonucleotides released from uniformly ^{32}P-labeled ϕX174 DNA by formic acid–diphenylamine hydrolysis. (d) Grid pattern of 3′,5′ terminally phosphorylated pyrimidine oligonucleotides derived from Figure 1c. (e) 5′-^{32}P terminally labeled pyrimidine oligonucleotides from ϕX174 DNA. (f) Grid pattern of 5′ terminally phosphorylated pyrimidine oligonucleotides derived from Figure 1e. (g) Uniformly ^{32}P-labeled terminally dephosphorylated pyrimidine oligonucleotides from ϕX174 DNA. (h) Grid pattern of terminally dephosphorylated pyrimidine oligonucleotides derived from Figure 1g.

Spleen and snake venom phosphodiesterase require a 5′- or 3′-hydroxyl terminus, respectively, on the oligonucleotide substrate, thus it was convenient to isolate the pyrimidine oligonucleotides of S13 and ϕX174 DNA in the dephosphorylated form prior to sequence analysis. Figure 1g shows the radioautogram and Figure 1h the corresponding grid pattern of terminally dephosphorylated ^{32}P-uniformly labeled pyrimidine oligonucleotides from ϕX174 DNA separated by ionophoresis–homochromatography. These data provided a second base composition assignment cross check. The short oligonucleotides are not present since they were removed by column chromatography to avoid overloading in

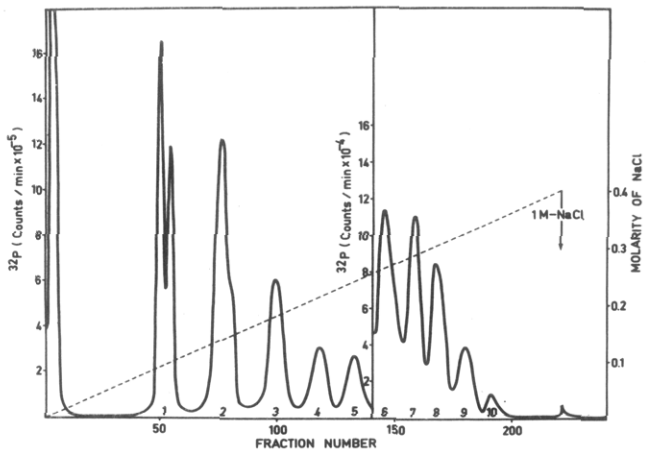

FIGURE 2: Chromatography of a formic acid–diphenylamine hydrolysate of uniformly ^{32}P-labeled ϕX174 DNA on DEAE-Sephadex A-25 according to chain length. For conditions see Methods and Černý et al. (1968, 1969). (—) ^{32}P radioactivity, (- - -) molarity of NaCl.

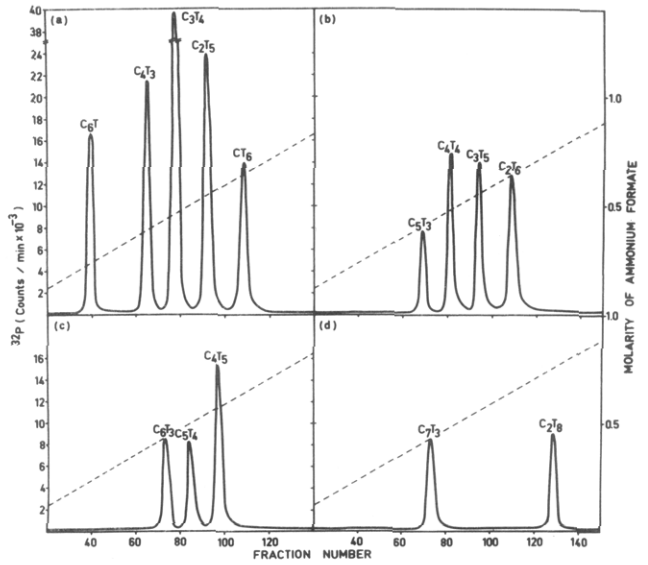

FIGURE 3: Chromatography of pyrimidine isostichs 7, 8, 9, and 10 of ϕX174 DNA on DEAE-Sephadex A-25 according to base composition. For conditions see Methods and Černý et al. (1968, 1969). (a) Isostich 7; (b) isostich 8; (c) isostich 9; (d) isostich 10. (—) ^{32}P radioactivity, (- - -) molarity of ammonium formate.

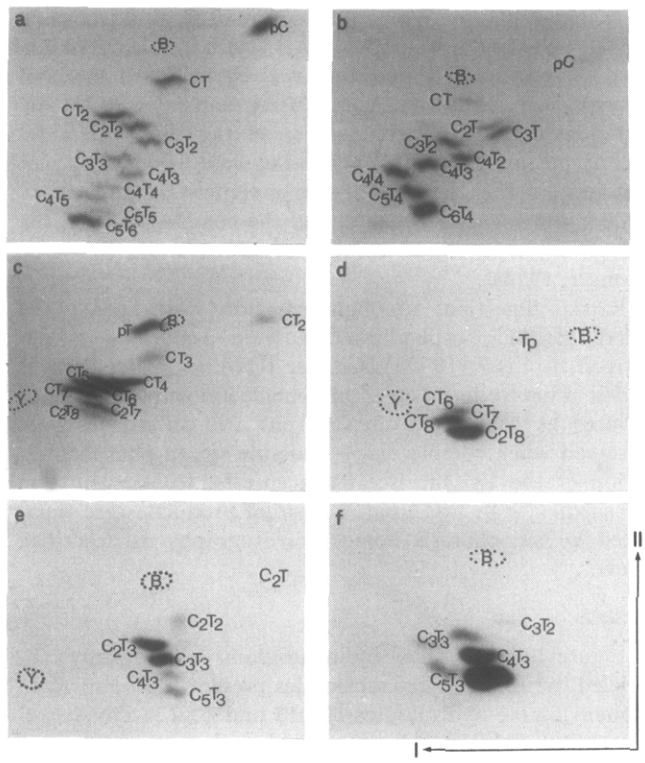

FIGURE 4: Two-dimensional fractionation of the products resulting from partial enzyme digestion of pyrimidine oligonucleotides from S13 DNA. (a) Oligonucleotide C_5T_6, 5′-^{32}P terminally labeled, digested with snake venom phosphodiesterase. (b) Oligonucleotide C_6T_4, 5′-^{32}P terminally labeled, digested with snake venom phosphodiesterase. (c) Oligonucleotide C_2T_8 uniformly ^{32}P-labeled, digested with snake venom phosphodiesterase. (d) Oligonucleotide C_2T_8 uniformly ^{32}P-labeled, digested with spleen phosphodiesterase. (e) Oligonucleotide C_5T_3 uniformly ^{32}P-labeled, digested with snake venom phosphodiesterase. (f) Oligonucleotide C_5T_3 uniformly ^{32}P-labeled, digested with spleen phosphodiesterase.

the ionophoresis step. It is obvious from Figure 1c, e, and g and the corresponding grid patterns Figure 1d, f, and h that the position of each oligonucleotide relative to the blue and yellow dye markers is dependent on the presence or absence of terminal phosphate groups. This difference is crucial in sequence analyses when base composition assignments are made of partial digestion products of dephosphorylated and 5′-phosphorylated oligonucleotides. To confirm the base composition assignments of the pyrimidine oligonucleotides of ϕX174 DNA presented in Figure 1d, a formic acid–diphenylamine hydrolysate of ϕX174 uniformly ^{32}P-labeled DNA was analyzed by column chromatographic methods (Černý et al., 1968, 1969). Figure 2 shows the chromatogram of the oligonucleotides separated according to chain length on a DEAE-Sephadex A-25 urea column. The longest oligonucleotide present was a decanucleotide; no undecanucleotides were detected. The same pattern was obtained in four different experiments on four different preparations of bacteriophage ϕX174 DNA. All isostichs were subfractionated according to base composition by chromatography on DEAE-Sephadex A-25 columns at pH 3.0. The elution profiles of isostichs 7, 8, 9, and 10 are shown in Figure 3. The distribution and base composition assignments of the pyrimidine oligonucleotides from ϕX174 DNA obtained by the column chromatographic procedures were identical with the distribution and assignments obtained by ionophoresis–homochromatography.

The long pyrimidine oligonucleotides from S13 and ϕX174 DNA isolated by both column fractionation techniques and ionophoresis–homochromatography were sequenced by partial digestion with snake venom and spleen phosphodiesterase and the resulting digestion products separated by ionophoresis–homochromatography. Figure 4a shows the separation of the partial snake venom phosphodiesterase products of 5′-^{32}P-labeled oligonucleotide C_5T_6 from S13 DNA. The separated products result from the progressive removal of nucleotides from the 3′-end of C_5T_6. Base compositions were assigned from the grid in Figure 1f. The first base removed was T, $C_5T_6 \rightarrow C_5T_5$ then C, $C_5T_5 \rightarrow C_4T_5$ then T, $C_4T_5 \rightarrow C_4T_4$, etc., giving the total sequence 5′-C-T-T-C-C-T-C-T-T-C-T-3′. Sequence analysis by partial snake venom phosphodiesterase digestion of 5′-^{32}P terminally labeled C_6T_4 from S13 DNA is shown in Figure 4b. The sequence is C-T-C-C-T-C-T-T-C-C. Figure 4c and d show separations of partial snake venom phosphodiesterase and spleen phosphodiesterase digests, respectively, of uniformly ^{32}P-labeled C_2T_8. In these experiments the

Table I: Sequences of Large Pyrimidine Oligonucleotides from S13 and φX174 DNA.[a]

Isostich	Composition	S13 DNA		φX174 DNA
11	C_5T_6	C-T-T-C-C-T-C-T-T-C-T		NP[b]
10	C_2T_8	C-T-T-T-T-T-T-T-C-T	⟷	C-T-T-T-T-T-T-T-C-T
	C_6T_4	C-T-C-C-T-C-T-T-C-C	⟷	NP
	C_7T_3	NP		C-T-C-C-T-C-T-C-C-C
9	C_4T_5	T-C-T-T-T-C-T-C-C	⟷	T-C-T-T-T-C-T-C-C
		C-T-C-T-T-T-C-T-C	⟷	C-T-C-T-T-T-C-T-C
	C_5T_4	C-T-T-C-C-T-C-C-T	⟷	C-T-T-C-C-T-C-C-T
	C_6T_3	T-C-C-T-T-C-C-C-C	⟷	T-C-C-T-T-C-C-C-C
8	C_2T_6	C-T-T-T-T-C-T-T	⟷	C-T-T-T-T-C-T-T
		T-T-T-T-C-T-T-C	⟷	T-T-T-T-C-T-T-C
	C_3T_5	3 molar yield		T-T-T-C-T-C-C-T
				T-C-T-T-C-T-T-C
	C_4T_4	C-T-T-C-C-T-T-C		C-T-T-C-C-T-T-C
				X
		NP		C-T-C-T-T-T-C-C
	C_5T_3	C-C-T-T-T-C-C-C	⟷	C-C-T-T-T-C-C-C
7	CT_6	C-T-T-T-T-T-T	⟷	C-T-T-T-T-T-T
		NP		T-T-T-C-T-T-T
	C_5T_2	C-T-C-T-C-C-C		NP
	C_6T	C-T-C-C-C-C-C	⟷	C-T-C-C-C-C-C
		NP		C-C-C-T-C-C-C

[a] Double-headed arrows indicate homology between the pyrimidine sequences in the DNA of two phages. Symmetrical sequences which read the same forwards and backwards are italicized and the center of symmetry indicated by a dot. [b] NP means that the oligonucleotide was not present. The sequences of the isomeric oligonucleotides C_4T_4 from φX174 DNA may be any of the possibilities indicated by the crossover position X. These oligonucleotides have not been included in any homology comparisons or calculations.

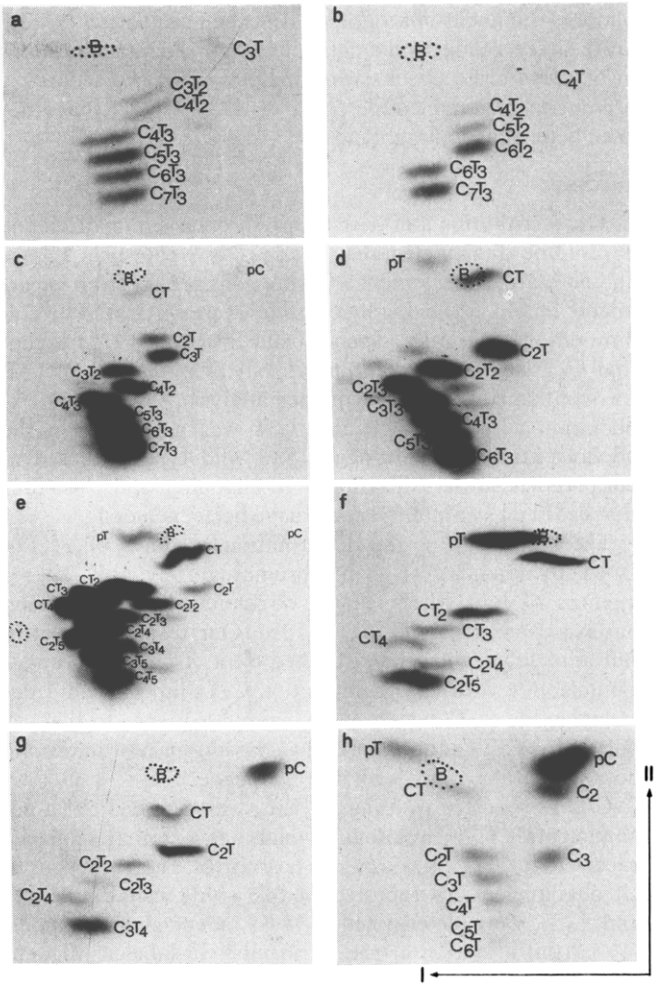

FIGURE 5: Two-dimensional fractionation of the products resulting from partial enzyme digestion of pyrimidine oligonucleotides from φX174 DNA. (a) Oligonucleotide C_7T_3 uniformly ^{32}P-labeled, digested with snake venom phosphodiesterase. (b) Oligonucleotide C_7T_3 uniformly ^{32}P-labeled, digested with spleen phosphodiesterase. (c) Oligonucleotide C_7T_3, 5′-^{32}P terminally labeled, digested with snake venom phosphodiesterase. (d) Oligonucleotide C_6T_3, 5′-^{32}P terminally labeled, digested with snake venom phosphodiesterase. (e) Oligonucleotide C_4T_5, 5′-^{32}P terminally labeled, digested with snake venom phosphodiesterase. (f) Oligonucleotide C_2T_5 from Figure 5e, digested with snake venom phosphodiesterase. (g) Oligonucleotide C_3T_4 from Figure 5e, digested with snake venom phosphodiesterase. (h) Oligonucleotide C_6T, 5′-^{32}P terminally labeled, digested with snake venom phosphodiesterase.

partial digestion products had no terminal phosphates and base compositions were assigned from the grid shown in Figure 1h. The snake venom phosphodiesterase (Figure 4c) products allow deduction of the sequence at the 3′-end of C_2T_8 as 5′---T-T-T-T-T-C-T. Similarly, by partial spleen phosphodiesterase digestion (Figure 4d), the sequence at the 5′-end has been deduced to be C-T-T-------3′. Combining the two results and the base composition C_2T_8 the total sequence can be constructed to be C-T-T-T-T-T-T-T-C-T. The radioautograms in Figure 4e and f show sequence analyses of uniformly ^{32}P-labeled S13 C_5T_3 by partial digestion with snake venom and spleen phosphodiesterases, respectively. The sequence at the 3′-end has been deduced to be 5′----T-T-C-C-C and the 5′-end C-C-T----3′, so that the total sequence is C-C-T-T-T-C-C-C.

The long pyrimidine oligonucleotides of φX174 DNA were sequenced in the same way as described for S13 DNA and some examples are shown in Figure 5. When uniformly ^{32}P-labeled C_7T_3 was digested with snake venom or spleen phosphodiesterase the separations shown in Figure 5a and b, respectively, were obtained. The sequence 5′----T-C-T-C-C-C at the 3′-end (Figure 5a) overlaps with the sequence C-T-C-C-T----3′ at the 5′-end (Figure 5b) to give the complete sequence C-T-C-C-T-C-T-C-C-C. The same sequence was obtained when 5′-^{32}P terminally labeled C_7T_3 was digested with snake venom phosphodiesterase (Figure 5c). Figure 5d shows the radioautogram of the fractionated partial digestion products obtained when 5′-^{32}P terminally labeled C_6T_3 was digested with snake venom phosphodiesterase. From the pattern the complete sequence can be deduced as T-C-C-T-T-C-C-C-C. Oligonucleotide C_4T_5 contains two sequences (Figure 5e). The first base removed from the 5′-^{32}P terminally labeled oligonucleotide by snake venom phosphodiesterase was a C. Removal of the second base resulted in two products C_2T_5 and C_3T_4. These two areas were scratched from the plate; the oligonucleotide material was eluted and again digested with snake venom phosphodiesterase (Figure 5f and g). Taken together the information obtained from Figure 5e, f, and g allows deduction of the two sequences T-C-T-T-T-C-T-C-C and C-T-C-T-T-T-C-T-C, present in oligonucleotide C_4T_5. Oligonucleotide C_6T also contains two sequences C-T-C-C-C-C-C and C-C-C-T-C-C-C (Figure 5h). In this case the two sequences were determined from one digest.

All the pyrimidine oligonucleotides from S13 and φX174 DNAs that are present in 1 or 2 *M* yield have been sequenced and are listed in Table I. Each oligonucleotide was isolated by both column fractionation methods and by iono-

phoresis–homochromatography and then sequenced by separate snake venom and spleen phosphodiesterase digestions in order to cross check every analysis. The two sequences present in oligonucleotide C_4T_4 of ϕX174 DNA have not been determined unambiguously.

Discussion

The distribution and base composition assignments of the pyrimidine oligonucleotides of S13 DNA shown in Figure 1a and b are in agreement with previous results from experiments in which the oligonucleotides were separated by column chromatography (Spencer and Boshkov, 1973) except that C_6T was not detected previously. The identity of C_6T was confirmed by direct sequence analysis. We have no explanation why oligonucleotide C_6T was not found in the previous study. Bacteriophage S13 wild type was used in both studies so that differences in oligonucleotide distribution based on strain differences have been excluded.

The separation of the depurination products of ϕX174 DNA by ionophoresis–homochromatography and the sequences of four of the larger oligonucleotides have been published previously (Ling, 1972b). Our results are not in full agreement with the published data. There is no undecanucleotide with base composition C_8T_3 but instead a decanucleotide C_7T_3 which has a similar sequence to that reported for C_8T_3 except for one C missing in position 3. No decanucleotide C_7T_3 with the sequence T-C-C-T-C-T-C-C-C-C as reported by Ling (1972b) was detected but a nonanucleotide C_6T_3 was found which has a similar sequence except that one C is missing in position 5. The sequences of all oligonucleotides reported in this study including C_7T_3 and C_6T_3 were determined by two independent methods: (1) partial digestion of the uniformly ^{32}P-labeled oligonucleotide with snake venom and spleen phosphodiesterase, and (2) partial digestion of the 5′-^{32}P terminally labeled oligonucleotide with snake venom phosphodiesterase. Both methods gave the same results. The distribution and base composition assignments of the pyrimidine oligonucleotides of ϕX174 DNA determined from ionophoresis–homochromatograms of uniformly ^{32}P-labeled, 5′-^{32}P terminally labeled, and uniformly ^{32}P-labeled terminally dephosphorylated hydrolysates were further supported by experiments in which the depurination products were separated by column chromatography (Figures 2 and 3). No undecanucleotide was detected, two decanucleotides C_7T_3 and C_2T_8 and three nonanucleotides C_6T_3, C_5T_4, and C_4T_5 were present. The mutant ϕX174*am*3 used in this study was the same as that used by Ling (1972b). However, the possibility that base mutations have occurred in stocks of ϕX174*am*3 maintained in different laboratories cannot be excluded at this time.

Considerable sequence homology was observed between the pyrimidine oligonucleotides of S13 and ϕX174 (Table I). Oligonucleotides C_6T and CT_6 of ϕX174 contain two isomeric sequences one of which is always identical with the unique sequence of the corresponding oligonucleotides in S13. This may also apply to the C_4T_4 oligonucleotides. Oligonucleotides C_6T_4 and C_5T_2 found in S13 are not present in ϕX174 but appear to be related to C_7T_3 and C_6T of ϕX174 differing by a single T → C change in the sequence. The sequence of undecanucleotide C_5T_6 is unique to S13 and seems not to be related to any sequence in ϕX174. Of the 14 sequences reported from S13 DNA (120 bases) and listed in Table I, at least ten are identical with sequences from ϕX174 DNA (92 bases) and two oligonucleotides are closely related (17 bases), the only difference being a T → C transition (a total of 107 identical bases).

The extensive sequence homology between ϕX174 and S13 DNA is not surprising since the two phages have been shown to be closely related by genetic, immunological, and chemical studies (for references see introduction). Recently the sequence relationship between ϕX174 and S13 DNA has been investigated by degradation with four different restriction enzymes (Hayashi and Hayashi, 1974; Grosveld et al., 1976). The studies support the close sequence relationship between the two phages. Differences occur in two regions of the genome only (Grosveld et al., (1976). Godson (1973) using heteroduplex formation showed that only one small segment (4.7 ± 1.9%) of the DNA of ϕX174 is highly homologous with S13 DNA; the rest is partially homologous with an overall average 36% base mismatch. Our results indicate a much closer sequence relationship between the two phages. However, we cannot exclude the possibility that long pyrimidine sequences may have been preserved during the evolution of related phages. Final conclusions on sequence homology must await more extensive sequence data from both phages.

The same degree of homology described for the pyrimidine oligonucleotide sequences of S13 and ϕX174 has also been observed between the closely related filamentous bacteriophages fd and fl (Ling, 1972b).

The relationship of the filamentous bacteriophages (fd, fl) with the spherical bacteriophages (ϕX174, S13) is less clear. The phages of both groups contain single-stranded circular DNAs of approximately the same molecular weight (Marvin and Hohn, 1969). The DNAs are thymine-rich and replicate via a double-stranded replicative form intermediate. When the longer pyrimidine oligonucleotide sequences of S13 and ϕX174 are compared with those of fd or fl (Ling, 1972b) partial homology is observed. For example, the sequence C-T-T-C-C-T-C-T-T is part of C_4T_5 of fd and fl and C_5T_6 of S13, the sequence C-T-T-T-T-T-T-T part of C_3T_8 of fd and fl and C_2T_8 of S13 and ϕX174, the sequence T-C-C-T-T-C part of C_4T_6 of fd and fl and C_6T_3 of S13 and ϕX174. However, none of the sequences of fd and fl are completely identical with any of the sequences described in S13 or ϕX174. The discovery of some sequence homology between the DNAs of the spherical and filamentous phages suggests a distant evolutionary relationship.

A comparison of pyrimidine sequences from S13, ϕX174, fd, and fl with those from bacteriophage λ (Hewish et al. in preparation) shows that all five phage DNAs have the following sequences in common: C-T-T-T-T-T-T-T, C-C-T-T-C-C, T-T-T-C-T-T, and T-T-C-C-T-T. The significance of this observation is not known. Bacteriophage λ and the small DNA phages may be distantly related. The common sequences may be part of a genome region which has the same function and the sequences have been evolutionarily preserved.

In Table II the pyrimidine oligonucleotide sequences of S13 DNA are listed to show internal sequence homology. In some oligonucleotides, e.g., C_5T_6 and C_6T_4, up to eight bases are homologous. The sequence T-C-C-T is shared by five different oligonucleotides and in two oligonucleotides C_4T_5 and C_2T_6 extensive homology exists between the isomeric sequences. The same internal homology is present in ϕX174 DNA (see Table I) and has been observed previously in fd DNA (Ling, 1972a). The significance of this observation is not clear. One explanation is that these sequences arose by amplification during the evolution of the phage

Table II: Homologous Sequences among the Pyrimidine Oligonucleotides of S13 DNA.[a]

Composition	Sequence
C_5T_6	C-T-T-C-C-T-C-T-T-C-T
C_6T_4	C-T-C-C-T-C-T-T-C-C
C_4T_4	C-T-T-C-C-T-T-C
C_5T_4	C-T-T-C-C-T-C-C-T
C_5T_3	C-C-T-T-T-C-C-C
C_6T_3	T-C-C-T-T-C-C-C-C
CT_6	C-T-T-T-T-T-T
C_2T_8	C-T-T-T-T-T-T-T-C-T
C_4T_5	T-C-T-T-T-C-T-C-C
C_4T_5	C-T-C-T-T-T-C-T-C
C_2T_6	C-T-T-T-T-C-T-T
C_2T_6	T-T-T-T-C-T-T-C
C_4T_4	C-T-T-C-C-T-T-C
C_5T_4	C-T-T-C-C-T-C-C-T
C_5T_6	C-T-T-C-C-T-C-T-T-C-T
C_6T_3	T-C-C-T-T-C-C-C-C
C_6T_4	C-T-C-C-T-C-T-T-C-C

[a] Boxed areas indicate homologous sequences.

DNA.

Another feature of the oligonucleotides listed in Table I is the occurrence of symmetrical sequences or true palindromes (italicized in Table I) which read the same forwards and backwards. Recently Pieczenik et al. (1974) found sequences with the same kind of symmetry in ribosome binding sites of fl RNA and pointed out that they are common to several DNAs and RNAs near initiation codons. In contrast to the self-complementary palindromic sequences present in double-stranded DNA (Wilson and Thomas, 1974) which can form hairpin structures around their axes of symmetry true palindromes are inherently incapable of such hydrogen-bonded structures and may act to prevent hairpin formation in defined areas of the genome.

At the present time it is not known if long pyrimidine oligonucleotides have any specific function. It was suggested by Szybalski et al. (1966) that cytosine-rich oligonucleotides are related to the initiation and termination regions for RNA transcription. These cytosine-rich sequences were inferred from the existence of sites in single-stranded DNA which bind poly(G). This received partial support from the observation that the long pyrimidine oligonucleotides of T7 DNA occur predominantly in the transcribed strand (Mushynski and Spencer, 1970). However, studies of the distribution of poly(G) binding sites in λ DNA (Champoux and Hogeness, 1972) and sequence studies of promotor regions (e.g., Maniatis et al., 1974) have shown no involvement of long pyrimidine oligonucleotides with initiation of transcription.

Infrared dichroism (Pilet and Brahms, 1972) and x-ray diffraction (Arnott and Selsing, 1974a,b; Arnott et al., 1974) have shown that base composition influences the conformation of a DNA molecule. High AT-content favors B or B-like conformations and for very AT-rich DNA duplexes the B → A transition is difficult. Arnott and Selsing (1974a,b) and Arnott et al. (1974) have shown that base sequence also influences the conformations that may be assumed by DNA double helices and that the effect of alternating purine/pyrimidine sequences is quite different from that of homopurine/homopyrimidine sequences. This property might enhance the distinctiveness of such stretches in DNAs such as S13, φX174, fd, and fl where these sequences may have a role as recognition sites or controllers of the rate of transcription (Dickson et al., 1975). Arnott and Bond (1973) and Arnott and Selsing (1974a) have suggested that homopurine/homopyrimidine blocks might function biologically through their ability to participate in triple-stranded complexes with an additional complementary oligopyrimidinenucleotide. The conformations in these triple-stranded structures are always A, and S. Arnott (personal communication) has pointed out that an oligopyrimidinenucleotide acting on homopurine/homopyrimidine sequences about ten nucleotide pairs long might not only induce A-like conformations local to the site of triplex formation but also prompt longer stretches of the DNA duplex to be A also. In this way these sequences could be involved in events such as transcription without being part of the promotor region.

Acknowledgments

The authors thank Dr. S. Eisenberg for assistance in the isolation of φX174 DNA, Mr. A. Kabassakalian for continued able technical assistance, and Dr. Struther Arnott for helpful comments which contributed significantly to the discussion.

References

Arnott, S., and Bond, P. J. (1973), *Nature (London), New Biol. 244,* 99–101.

Arnott, S., Chandrasekaran, R., Hukins, D. W. L., Smith, P. J. C., and Watts, L. (1974), *J. Mol. Biol. 88,* 523–533.

Arnott, S., and Selsing, F. (1974a), *J. Mol. Biol. 88,* 509–521.

Arnott, S., and Selsing, F. (1974b), *J. Mol. Biol. 88,* 551–552.

Baker, A., and Tessman, E. (1967), *Proc. Natl. Acad. Sci. U.S.A. 58,* 1439–1445.

Barrell, B. G., Weith, H. L., Donelson, J. E., and Robertson, H. D. (1975), *J. Mol. Biol. 92,* 377–393.

Benbow, R. M., Hutchison, C. A., III, Fabricant, J. D., and Sinsheimer, R. L. (1971), *J. Virol. 7,* 549–558.

Bernardi, A., and Bernardi, G. (1968), *Biochim. Biophys. Acta 155,* 360–370.

Brownlee, G. G., and Sanger, F. (1969), *Eur. J. Biochem. 11,* 395–399.

Burton, K. (1967), *Methods Enzymol. 12A,* 222–224.

Černý, R., Černá, E., and Spencer, J. H. (1969), *J. Mol. Biol. 46,* 145–156.

Černý, R., Mushynski, W. E., and Spencer, J. H. (1968), *Biochim. Biophys. Acta 169,* 439–450.

Champoux, J., and Hogeness, D. S. (1972), *J. Mol. Biol. 71,* 383–405.

Denhardt, D. T. (1969), *J. Biol. Chem. 244,* 2710–2715.

Denhardt, D. T., Dressler, D. H., and Hathaway, A. (1967), *Proc. Natl. Acad. Sci. U.S.A. 57,* 813–820.

Dickson, R. C., Abelson, J., Barnes, W. M., and Reznikoff, W. S. (1975), *Science 187,* 27–35.

Fiers, W., and Sinsheimer, R. L. (1962), *J. Mol. Biol. 5,* 424–428.

Galibert, F., Sedat, J., and Ziff, E. (1974), *J. Mol. Biol. 87,* 377–407.

Garen, A., and Levinthal, C. (1960), *Biochim. Biophys. Acta 38,* 470–483.

Godson, N. G. (1973), *J. Mol. Biol. 77,* 467–477.

Grosveld, F. G., Ojamaa, K. M., and Spencer, J. H. (1976),

Virology (in press).
Hänggi, U. J., Streeck, R. E., Voigt, H. P., and Zachau, H. G. (1970), *Biochim. Biophys. Acta 217,* 278–293.
Harbers, K., Harbers, B., and Spencer, J. H. (1974), *Biochem. Biophys. Res. Commun. 58,* 814–821.
Harbers, K., and Spencer, J. H. (1974), *Biochemistry 13,* 1094–1101.
Hayashi, M. N., and Hayashi, M. (1974), *J. Virol. 14,* 1142–1151.
Jay, E., Bambara, R., Padmanabhan, R., and Wu, R. (1974), *Nucleic Acids Res. 1,* 331–353.
Jeng, Y., Gelfand, M., Hayashi, M., Schleser, R., and Tessman, E. S. (1970), *J. Mol. Biol. 49,* 521–526.
Ling, V. (1972a), *J. Mol. Biol. 64,* 87–102.
Ling, V. (1972b), *Proc. Natl. Acad. Sci. U.S.A. 69,* 742–746.
Maniatis, T., Ptashne, M., Barrell, B. G., and Donelson, J. (1974), *Nature (London) 250,* 394–397.
Marvin, D., and Hohn, B. (1969), *Bacteriol. Rev. 33,* 172–209.
Mushynski, W. E., and Spencer, J. H. (1970), *J. Mol. Biol. 52,* 91–106.
Pieczenik, G., Model, P., and Robertson, H. D. (1974), *J. Mol. Biol. 90,* 191–214.
Pilet, J., and Brahms, J. (1972), *Nature (London), New Biol. 236,* 99–100.
Richardson, C. C. (1965), *Proc. Natl. Acad. Sci. U.S.A. 54,* 158–165.
Schendel, P. F., and Wells, R. D. (1973), *J. Biol. Chem. 248,* 8319–8321.
Sinsheimer, R. L. (1959), *J. Mol. Biol. 1,* 43–53.
Sinsheimer, R. L. (1966), *Proced. Nucleic Acid. Res. 1,* 569–576.
Spencer, J. H., and Boshkov, L. (1973), *Can. J. Biochem. 51,* 1206–1211.
Spencer, J. H., Černý, R., Černá, E., and Delaney, A. D. (1972), *J. Virol. 10,* 134–141.
Sulkowski, E., and Laskowski, M., Sr. (1971), *Biochim. Biophys. Acta 240,* 443–447.
Szybalski, W., Kubinski, H., and Sheldrick, P. (1966), *Cold Spring Harbor Symp. Quant. Biol. 31,* 123–127.
Tessman, I. (1959), *Virology 7,* 263–275.
Tessman, E. S. (1965), *Virology 25,* 303–321.
Tessman, E. S., and Schleser, R. (1963), *Virology 19,* 239–240.
Trim, A. R., and Dickerson, P. E. (1974), *Anal. Biochem. 57,* 93–99.
Warner, H. R., and Lewis, N. (1966), *Virology 29,* 172–175.
Wilson, D. A., and Thomas, C. A. (1974), *J. Mol. Biol. 84,* 115–144.
Wu, R., and Kaiser, A. D. (1967), *Proc. Natl. Acad. Sci. U.S.A. 57,* 170–177.
Yamamoto, K., Alberts, B. M., Benzinger, R., Lawhorne, L., and Treiber, G. (1970), *Virology 40,* 734–744.
Zahler, S. A. (1958), *J. Bacteriol. 75,* 310–315.

Phosphorylation and DNA Binding of Nuclear Rat Liver Proteins Soluble at Low Ionic Strength†

Archie W. Prestayko,* Paula M. Crane, and Harris Busch

ABSTRACT: Proteins were extracted from isolated rat liver nuclei with 0.15 M NaCl and 0.35 M NaCl at pH 8.0. The number of phosphoproteins in these extracts was determined by labeling with ^{32}P and autoradiography after two-dimensional gel electrophoresis. Two proteins, B22p and B24p, contained small amounts of ^{32}P and sedimented with the 30S nuclear informofer particle. With the exception of two phosphoproteins, CB and CN′, all of the phosphoproteins found in the 0.35 M NaCl extract of nuclei were also present in the 0.15 M NaCl extract. Approximately 20% of the 0.15 M NaCl soluble proteins bound to rat liver DNA in 0.05 M KCl–0.05 M Tris-HCl (pH 8). Of these proteins, 1–2% bound to DNA in 0.15 M KCl and were eluted with 2 M KCl. This DNA bound fraction which contained both phosphorylated and nonphosphorylated proteins was similar in both the 0.15 and 0.35 M NaCl extracts. However, two major proteins (C13 and C14) and three minor proteins (C15, C25, Cg′) were present only in the 0.15 M NaCl extract. The results of the present study show that there are marked similarities in the two-dimensional gel electrophoretic, phosphorylation, and DNA binding properties of rat liver nuclear proteins soluble in either 0.15 or 0.35 M NaCl.

Recently, studies have been made on nuclear proteins and their role in chromatin structure and function, particularly proteins "loosely" bound to chromatin that are extracted at low ionic strength (Patel, 1972; Comings and Tack, 1973; Holoubek and Fujitani, 1973; Kostraba and Wang, 1973; Kostraba et al., 1975). The involvement of "tightly" bound proteins in chromatin structure and function (Paul and Gilmour, 1966, 1968; Spelsberg and Hnilica, 1969; Spelsberg et al., 1971; Kostraba and Wang, 1972) has been recently reviewed (Olson and Busch, 1974; Stein et al., 1974).

Early studies on proteins soluble in Tris-saline buffers (Gurdon and Brown, 1965; Goldstein and Prescott, 1967, 1968; Gurdon, 1970; Goldstein, 1974) in *Amoeba* and *Xenopus* demonstrated that a number of nuclear proteins diffuse readily across the nuclear membrane; they have been referred to as cytonucleoplasmic shuttling proteins. More

† From the Department of Pharmacology, Baylor College of Medicine, Houston, Texas 77025. *Received August 7, 1975.* These studies were supported by the Cancer Research Center Grant CA-10893, P .5 from the National Cancer Institute.