Distribution of Pyrimidine Sequences in Bacteriophage TP-84 Deoxyribonucleic Acid†

Paul E. Kizer and Grady F. Saunders*

ABSTRACT: The distribution of A·T and G·C base pairs along the bacteriophage TP-84 DNA molecule was investigated. A graph of the differential of the thermal denaturation profile of TP-84 DNA resulted in three distinct compositional regions each with a characteristic T_m. By constructing a Felsenfeld denaturation graph the TP-84 genome showed at least one region comprising approximately 15% of the molecule which appears to be rich in dG·dC pairs. Most of the molecule (85%) has a base composition analogous to a uniform distribution of A·T and G·C base pairs. The pyrimidine isostich catalog showed a distribution of total clusters similar to that previously published for λ DNA (Mushynski, W. E., and Spencer, J. H. (1970b), *J. Mol. Biol. 52*, 107). λ DNA and TP-84 DNA have 5 runs/genome each of clusters 12 and 13 while T7 DNA has 7 runs in cluster 13 (Mushynski, W. E., and Spencer, J. H. (1970a), *J. Mol. Biol. 52*, 91). A total of 68 nonisomeric components was found with no deoxycytidylate runs longer than 6 nucleotides and no thymidylate runs longer than 9 nucleotides. T_9 is the longest thymidylate run to be reported so far and T_8 is absent. The C_6 and T_9 moieties were found in amounts corresponding to one run per TP-84 DNA molecule each.

The regulation of gene activity is due, in part, to the ability of certain proteins to recognize specific base sequences in DNA. The β-galactosidase (Gilbert and Muller-Hill, 1966) and bacteriophage λ repressors (Ptashne, 1967; Steinberg and Ptashne, 1971) of *Escherichia coli* have been shown to bind to DNA containing the appropriate regulatory genes while not binding to DNA of the corresponding deletion mutants. The DNA sequence recognized by DNA-dependent RNA polymerase for the binding and initiation of RNA transcription has not been established. However, studies with certain bacillus bacteriophages showed that the RNA synthesized during phage infection hybridizes exclusively with DNA strands bearing deoxycytidylate clusters (indicated by formation of stable DNA·poly(G) complexes) (Kubinski *et al.*, 1966). This evidence has been taken to implicate dC clusters in the DNA strand selection process (Szybalski *et al.*, 1968). Presumably, RNA polymerase can recognize either a sufficiently long dC sequence or the slight modification in DNA conformation resulting from a dC:dG stretch.

Infection of the appropriate host with bacteriophage TP-84 or T7 results in the synthesis of phage-specific RNA which hybridizes exclusively with one of the DNA strands, the strand containing dC clusters (Shaw, 1969; Summers and Szybalski, 1968a). If these clusters can serve as recognition sequences for RNA polymerase, then relevant information could be obtained from a pyrimidine tract analysis of the DNA molecule.

The pyrimidine tract analyses of bacteriophage DNAs published to date (ΦX-174: Darby *et al.*, 1970; S 13: Cerny *et al.*, 1969; T 7, λ: Mushynski and Spencer, 1970a,b) showed no cytosine runs longer than six residues. From these data Mushynski and Spencer (1970a) postulated the binding sites to be rich in T as well as dC.

Previous studies have shown that one strand (H strand) of TP-84 DNA hybridizes with phage-induced RNA and forms stable complexes with poly(rG), poly(I,G), and polyxanthylate (Saunders *et al.*, 1969). The complementary strand (L strand) complexes with poly(U). The present study was designed to show the frequency distribution of specific pyrimidine runs (clusters, isostichs; Shapiro *et al.*, 1965) in TP-84 DNA.

Materials and Methods

Bacteria and Bacteriophage Strains. Bacillus stearothermophilus strain 10 was used as the host for the growth of bacteriophage TP-84. This phage is specific for certain strains of *B. stearothermophilus*, an obligate thermophilic bacterium, and is routinely grown at 55°.

Growth Media and Reagents. The growth medium used for phage production, TCG medium, was low in phosphate and contained the following: 0.1 M Tris-HCl (pH 7.4), 8.5×10^{-3} M NaCl, 1.6×10^{-4} M Na_2SO_4, 3.2×10^{-4} M KH_2PO_4, 1×10^{-3} M $MgSO_4$, 1.4×10^{-3} M $CaCl_2$, 3×10^{-6} M $FeCl_3$, 2×10^{-5} M $MnCl_2$, 10 ml/l. of 5% casein hydrolysate (vitamin free, enzymatic from Nutritional Biochemicals Corp.), 5 ml/l. of 20% glucose, and 5 ml/l. of vitamin solution containing 15 mg/l. of thiamine, 15 mg/l. of riboflavin, 150 mg/l. of nicotinic acid, 100 mg/l. of pantothenic acid, 7.5 mg/l. of pyridoxine, and 900 mg/l. of biotin.

Stock solutions of 2 M triethylammonium bicarbonate (TEAB) were prepared by mixing appropriate amounts of triethylamine with H_2O and sparging with CO_2 in an ice bath with constant stirring until the pH reached 8.5.

Carrier-free ^{32}P in the H_3PO_4 form was obtained from International Chemical and Nuclear Corp. Standard saline citrate, SSC, is 0.15 M NaCl and 0.015 M trisodium citrate (pH 7.0). Alkaline phosphatase (ribonuclease free), DNase I (RNase free), and snake venom phosphodiesterase were obtained from Worthington Biochemical Corp. *p*-Nitrophenyl phosphate reagent for the alkaline phosphatase test and calf thymus DNA, sodium salt grade A, were obtained from

† From the Department of Developmental Therapeutics, The University of Texas at Houston, M. D. Anderson Hospital and Tumor Institute, Houston, Texas 77025, and The University of Texas Graduate School of Biomedical Sciences, Houston, Texas 77025. *Received September 9, 1971.* This investigation has been supported by Grant G-267 from the Robert A. Welch Foundation. Part of the thesis of P. E. K., presented to The University of Texas at Houston Graduate School of Biomedical Sciences in partial fulfillment of requirements for the M. S. degree.

Calbiochem. Diethylaminoethyl (DEAE)-cellulose was obtained from Bio-Rad Laboratories under the brand name Cellex-D with an exchange capacity of 0.6 mequiv/g.

Phage Growth and Purification. Four 500-ml Delong flasks each containing 250 ml of TCG media were inoculated with 15 ml of an overnight bacterial culture. The flasks were shaken at 55° until the absorbance at 420 nm of the cultures reached 0.2 as measured on a Bausch and Lomb Spectronic 20 and the cells collected by a 5-min centrifugation at 8000*g*. The cells were resuspended in TCG medium, glucose added to 0.5%, ^{32}P added to give a final concentration of 200 μCi/ml, and phage added to give 5×10^8 pfu/ml. Lysis was usually complete after 4 hr of shaking at 55°. NaCl was added to the lysate to 0.5 M and the flasks were permitted to shake while cooling to room temperature. The cell debris was removed by centrifugation at 8000*g* for 10 min. Poly(ethylene glycol) (mol wt 6000–7500) was added to the supernatant fraction to 6% w/v. After standing in the cold for at least 1 hr the phage particles were collected by a 10-min centrifugation at 7000*g* and resuspended in 4 ml of SSC. The phage were purified by banding in CsCl density gradients as previously described (Saunders and Campbell, 1965).

Preparation of Phage DNA. The purified phages were dialyzed against SSC for at least 6 hr at 4° and the DNA extracted by shaking 15 min at 4° with an equal volume of redistilled phenol saturated with 0.1 M Tris-HCl (pH 6.5). The aqueous phase was removed, extracted with diethyl ether, and the residual ether removed by sparging with N_2. The purity of the phage DNA preparations is routinely monitored by banding in CsCl density gradients (Saunders and Campbell, 1965). Host bacterial DNA which is 9% higher in G + C content was not detected in our phage DNA preparations.

Preparation of Pyrimidine Oligonucleotides. Calf thymus carrier DNA (250 mg) was added as carrier to the ^{32}P phage DNA (~100 μg) solution. DNA degradation with diphenylamine and formic acid was carried out as described by Burton and Peterson (1960). After addition of one-half volume of water the DNA digest was extracted three times with six volumes each of ether. Residual ether was eliminated with N_2, formic acid and water were removed by lyophilization, and the residue dissolved in 7 M urea–0.1 M acetate buffer (pH 5.5).

Separation of Pyrimidine Isostichs. Fractionation of pyrimidine oligonucleotides according to chain length was carried out as described by Peterson and Reeves (1966). A DEAE-cellulose column (1.8 × 40 cm), equilibrated with 7 M urea–acetate buffer (pH 5.5), was loaded with the DNA digests. Fractionation was accomplished by elution with two linear NaCl gradients in urea–acetate buffer, the first from 0.0 to 0.2 M and the second from 0.17 to 0.30 M. The absorbance at 254 nm was continuously monitored and 4-ml fractions were collected. This wavelength is close to the 249-nm isosbestic wavelength at pH 5 of pyrimidine deoxynucleotides (Spencer *et al.*, 1968). The ^{32}P Cerenkov radiation (Clausen, 1968) was measured by counting the total volume of each fraction in the tritium channel of a Packard Tri-Carb Model 3375 liquid scintillation spectrometer.

Separation of Compositional Isomers. Peak fractions of each pyrimidine isostich were pooled, diluted threefold with distilled water, and applied to a 1 × 20 cm DEAE-cellulose column. Following loading, the column was washed with a 7 M urea–0.1 M formic acid buffer (pH 3.5). The isomers were eluted with linear NaCl gradients (which varied slightly according to chain length) in urea–formate buffer (Rushizky and Sober, 1964). Fractions were collected and absorbancy and radioactivity monitored as described above.

Desalting the Oligonucleotides. The peak fractions were pooled, diluted threefold, and readsorbed into a 1 × 10 cm DEAE-cellulose column. Each column was then washed with 0.01 M TEAB buffer to remove the urea and salts (Peterson and Reeves, 1966). Oligonucleotides were eluted with 2 M TEAB, a volatile salt, which was subsequently removed by lyophilization.

Determination of Chain Length. The desalted oligonucleotides dissolved in 4 ml of 4 M NaCl–0.1 M Tris-HCl (pH 8.0) were incubated with alkaline phosphatase (25 units/ml) for 30 min at 60°. The high-salt concentration was used to inhibit nuclease activity. The oligonucleotides were diluted with H_2O, adsorbed to a 1 × 20 cm DEAE-cellulose column, and eluted with a NaCl gradient. Inorganic phosphate eluted during column loading and the remaining material was eluted with a 0.0–0.5 M NaCl gradient. Alkaline phosphatase activity was identified using a *p*-nitrophenyl phosphate reagent. The ^{32}P counts associated with the enzyme activity were considered as inorganic phosphate. The nucleotide-protected phosphate elutes last in the gradient. Fractions were collected and uv absorbance and radioactivity were monitored as described above.

Composition of Oligonucleotides. The compositions of the oligonucleotides were determined by digesting the desalted components with 50 μg/ml of DNase I (10.5 units) in 0.005 M $MgCl_2$–0.005 M Tris-HCl (pH 7.0) for 3 hr at 37°. After boiling for 5 min the pH was adjusted to 9 with NaOH, and glycine buffer was added to 0.02 M. Digestion to the mononucleotide level was accomplished by incubation for 3 hr at 37° with 0.01 unit of venom phosphodiesterase.

Some of the samples were applied to Dowex-1 formate columns for analysis by the method of Canellakis and Mantsavinos (1957). Most of the samples of mononucleotides were digested with alkaline phosphatase and applied to a nucleoside analyzer using a Bio-Rad A-6 column according to Uziel *et al.* (1968).

Denaturation Analyses. TP-84 DNA was dialyzed against either 0.1 × SSC, for the differential melting curve, or 0.01 M NaCl–0.001 M phosphate (pH 7.0), for the denaturation graph. Thermal denaturation experiments were carried out in a Zeiss PMQ II recording spectrophotometer equipped with an automatic sample changer. Cuvet temperature, controlled by a Lauda P120 linear electronic programmer, was monitored with a Gilford 207 linear thermosensor inside a blank cuvet, and recorded by a Texas Instrument Multi-Riter six-channel recorder. The temperature was increased automatically at 0.2°/min during the differential T_m and absorbances during the thermal transition were recorded continuously.

For the denaturation graph, the temperature was increased manually in 0.5° increments and allowed to equilibrate for at least 15 min before each reading. Readings were made at 5-nm intervals from 220 to 290 nm. All absorbance measurements were corrected for thermal expansion of water. The data were analyzed by the methods described by Felsenfeld and coworkers (Felsenfeld and Sandeen, 1962; Felsenfeld and Hirschman, 1965; Hirschman and Felsenfeld, 1966; Hirschman *et al.*, 1967) using their parameters for a quadratic analysis. The purified TP-84 DNA used in these analyses had uv absorbance ratios $A_{280}/A_{260} = 0.532$ and $A_{230}/A_{260} = 0.423$.

Results

The presence of regions in DNA with substantially differing base compositions can be detected by plotting the DNA melting curve in differential form. This differential T_m is ob-

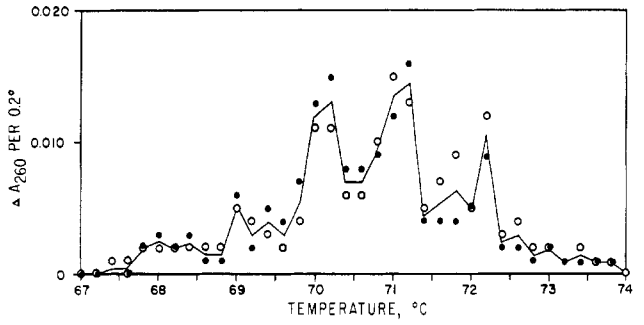

FIGURE 1: Differential melting curves of TP-84 DNA in 0.0195 M Na^+ (0.1 × SSC). The same DNA preparation was used in duplicate experiments (open and closed circles). Cuvet temperatures were measured directly by means of a thermocouple. The change in absorbance per 0.2° is plotted on the ordinate and temperature is on the abscissa.

tained by plotting the absorbance increment per 0.2° *vs.* the actual temperature at which the absorbance change occurs. Figure 1 shows the results of two separate analyses of the differential melting curve of TP-84 DNA. The graph shows two major regions of melting with T_m's of ~70 and ~71° and what may be a third region at 72°. The rest of the transition is represented by the small increments in A_{260} at the first and last parts of the graph.

A second method for showing compositional differences along the molecule was developed by Felsenfeld and Hirschman (1965). With this method the distribution of (A + T)- and (G + C)-rich regions in DNA can be determined from a careful spectral analysis of the thermal denaturation process. At a number of temperatures during the thermal transition, the absorbance of highly purified DNA was measured at 5-nm increments from 220 to 290 nm. The data from these sequential spectra were then analyzed according to Hirschman and Felsenfeld (1966). The points along the line are generated as a fraction of the genome denatured. The fraction of total A·T base pairs denatured is on the ordinate with the fraction of G·C pairs denatured on the abscissa. The position of the points along the line represents the relative quantities of each type denaturing simultaneously. The denaturation graph of TP-84 DNA is shown in Figure 2. According to this analysis, approximately 85% of the genome is composed of regions in which A·T and G·C pairs are uniformly distributed and about 15% which is rich in G·C pairs. This late-denaturing (G + C)-rich region might correlate with the peak observed in Figure 1 at 72°. The denaturation graph was repeated five times and the figure shown is a good representative. Thus three distinct compositional regions are indicated by the differential Tm and at least two are suggested by the denaturation graph. The base composition in mole per cent of TP-84 DNA as determined with the nucleoside analyzer was G = 23.3, A = 29.5, C = 22.5, and T = 24.7. This compares to our previous reported values of G = 22.0, A = 27.7, C = 22.1, and T = 28.2 (Saunders and Campbell, 1965).

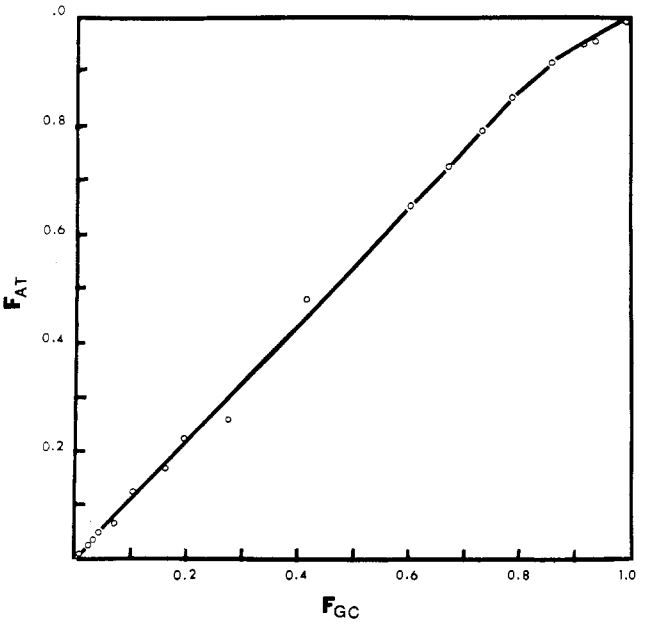

FIGURE 2: Denaturation graph of TP-84 DNA according to Hirschman and Felsenfeld (1966). $F_{A\cdot T}$ is the fraction of A·T pairs denatured and $F_{G\cdot C}$ is the fraction of G·C pairs denatured at each point in the course of denaturation. The solvent was 0.01 M NaCl–0.001 M Na_2PO_4 (pH 7.0).

TABLE I: Distribution of Pyrimidine Isostichs in TP-84.

Peak No.	Moles of Pyrimidine/ 100 g-atoms of DNA-P[a]	Number of Tracts/DNA Molecule[b]
1	16.71 ± 0.26	15,707
2	13.51 ± 0.09	6,350
3	7.81 ± 0.29	2,447
4	4.33 ± 0.04	1,018
5	2.69 ± 0.04	506
6	1.86 ± 0.13	291
7	1.34 ± 0.08	180
8	0.779 ± 0.048	92
9	0.429 ± 0.014	45
10	0.278 ± 0.023	26
11	0.127 ± 0.010	11
12	0.066 ± 0.001	5
13	0.070 ± 0.008	5
Total	49.999	

[a] Mean values for three separate analyses ±SEM and assumes pyrimidine accounts for 50% DNA-P. All samples were counted to <1% counting and the number of significant figures were determined accordingly. [b] Number of tracts calculated using TP-84 DNA of mol wt 32 × 10^6; 47,000 base pairs.

To directly examine DNA for runs of dC and T, pyrimidine tract analyses have proven useful (Spencer and Chargaff, 1963; Hall and Sinsheimer, 1963; Shapiro and Chargaff, 1964). [^{32}P]TP-84 DNA was mixed with enough calf thymus carrier DNA to allow the absorbance to be monitored and to provide an internal control. This mixture was digested to pyrimidine oligonucleotides with formic acid and diphenylamine and the digestion products separated according to chain lengths on a DEAE-cellulose–urea column. A typical separation is shown in Figure 3. Free purines and mononucleotides were eluted during the column loading and the 7 M urea–acetate wash, respectively. A clear separation of 13 isostichs corresponding to oligonucleotides of chain lengths from 1 to 13 nucleotides was obtained. Occasionally isostichs 14 and 15

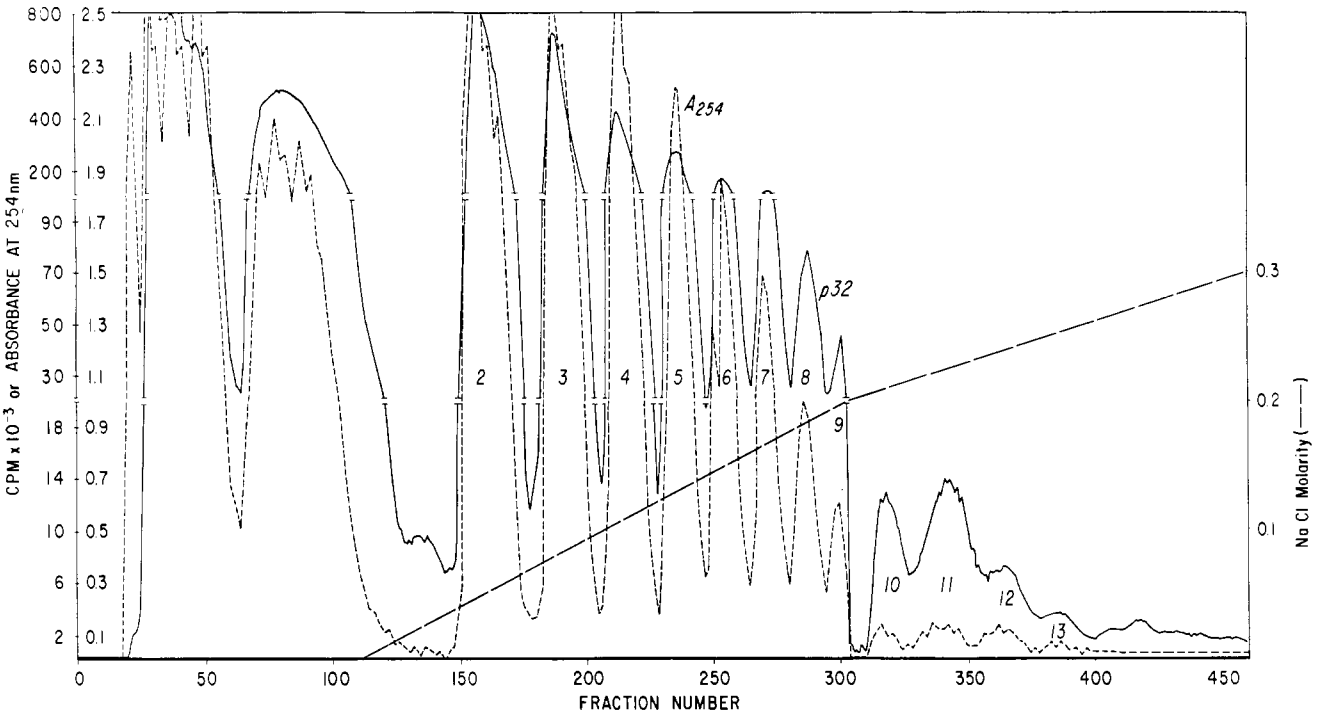

FIGURE 3: Separation of pyrimidine oligonucleotides into isostichs. Chromatography of a formic acid–diphenylamine hydrolysate of ^{32}P-labeled bacteriophage TP-84 DNA and carrier calf thymus DNA on a column of DEAE-cellulose (6 g, 0.69 mequiv/g, 35 × 1.5 cm chloride form). The hydrolysate (^{32}P-labeled TP-84 DNA, 9 × 10^7 cpm, 250 mg of calf thymus DNA carrier is 2.8 volumes of 2.7% diphenylamine in 90% formic acid, 37°, 18 hr) freed of formic acid and diphenylamine was applied to the column. The exchanger was washed with 0.1 M sodium acetate buffer in 7 M urea (pH 5.5) and then eluted with a linear gradient of NaCl to 0.2 M in 0.1 M sodium acetate buffer in 7 M urea (pH 5.5), followed by a second gradient of NaCl (0.17–0.30 M) in the same buffer. Total volume of eluent (1.6 l.); 4-ml fractions collected. The absorbance of the column effluent was monitored with an Isco UA2 scanner. Fractions were transferred to scintillation vials and Cerenkov radiation measured directly. (——) ^{32}P radioactivity; (– – –) absorbance at 254 nm; (— —) NaCl molarity.

were observed, but their specific activities were too low to allow further analyses. The first linear NaCl gradient elutes chains up to the nonanucleotide level and the second linear gradient elutes the remaining four to six isostichs. Following the second gradient a 0.5 M NaCl wash elutes some long-chain oligonucleotide material. After digestion to mononucleotides and analysis on a Varian nucleotide analyzer this fraction was found to contain nearly 10 mole % each of adenine and guanine. This was concluded to be a product of incomplete digestion. The quantitation of each isostich calculated by summation of the counts under each peak is shown in Table I. These data relate only to DNA phosphorus and, although quite unlikely, may not reflect the absolute amounts of pyrimidines. Lengths 12 and 13 have the same number of tracts per genome. The radioactivity recovered from the column (Figure 3), corrected for ^{32}P decay, was 68%. The high counts in the region of elution of inorganic phosphate (fractions 25–60) complicated accurate calculations of recovery.

The chain lengths of the isostichs were verified by three methods. In the first, the chain length is inferred from the elution position in the isostich separation. The second method involved determining the number of components that can be obtained from an isostich of short chain length. For example, a trinucleotide should have four possible components: C_3, C_2T, CT_2, and T_3 (see below). The third method consisted of assaying the relative number of phosphate groups cleaved by alkaline phosphatase digestion.

An isostich composed of a random distribution of C and T has the possibility of N + 1 compositional components, where N is the chain length in nucleotides. In order to more exactly determine the composition of each isostich, rechromatography of the isostichs at acid pH was undertaken. The pooled fractions from each chain length were diluted threefold and readsorbed to a fresh DEAE-cellulose column for elution with NaCl in 7 M urea buffered at pH 3.5 with 0.1 M formic acid. The NaCl gradient, different for each isostich, was adjusted according to the isostich chain length to ensure optimal separation. An example of the elution pattern for the components of some of the thirteen isostichs found in TP-84 DNA is shown in Figures 4–8.

The catalog of pyrimidine tracts in TP-84 DNA is given in Table II. Except for the skewness toward the thymidine-rich components, the amount of each component in TP-84 DNA generally followed a normal distribution in all isostichs in agreement with similar studies by other workers, *i.e.*, Rudner, *et al.*, 1966; Darby *et al.*, 1970; Cerny *et al.*, 1969. The closer to equimolar quantities of dC and T, the more total component is present. Exceptions noted were found in isostichs 8, 12, and 13. In number 8 the component C_4T_4 appears to be present in smaller quantities than C_5T_3. The C_5T_7 component of isostich 12 is missing, while C_4T_8 is present. C_6T_7 and C_4T_9 seem to be present in greater quantities than the other three components of number 13 but due to the chain length of the components the differences are not significant (see below).

Following summation of the counts under each peak, the quantitation of each component was calculated and the results are shown in Table II. Component C_6 is the longest cytosine run found in TP-84 DNA and it is present as a unique pyrimidine sequence. Of those oligonucleotides having a chain length of 9 or greater none contained components with only one thymidylate per chain. The component T_8 is missing in the octanucleotides but T_9 is present in isostich 9. This is the

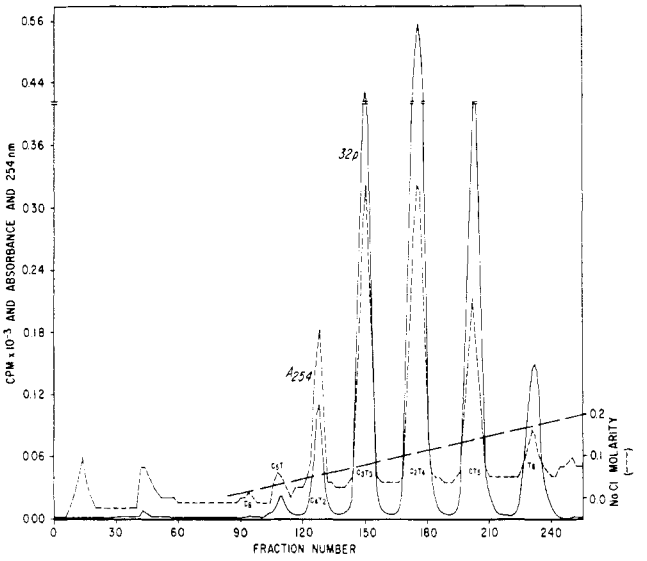

FIGURE 4: Fractionation of isostich 6. Rechromatography of pyrimidine isostichs from ^{32}P-labeled TP-84 DNA plus carrier calf thymus DNA on a column of DEAE-cellulose (0.69 mequiv/g, 20 × 1 cm, chloride form) at acid pH. The pooled isostich fractions were diluted threefold with water, and applied to the column. The column was washed with 0.1 M formic acid buffer (pH 3.5) in 7 M urea and elution effected with a linear gradient of NaCl in 0.1 M formic acid buffer (pH 3.5). Total volume of eluent 800 ml; 4-ml fractions collected. The recovery from this column was 93% of the input. (——) ^{32}P radioactivity; (– – – –) absorbance at 254 nm; (— —) NaCl molarity.

longest thymidylate isostich found and is also a unique pyrimidine sequence in TP-84 DNA. In isostich 13 a component with 0.009 mole of pyrimidine and one with 0.019 mole of pyrimidine are assigned the same number of tracts. The calculations which yield the number of tracts were all rounded to the next even number. In the example above 0.009 mole of pyrimidine corresponds to 0.650 tract/genome, while 0.019 mole

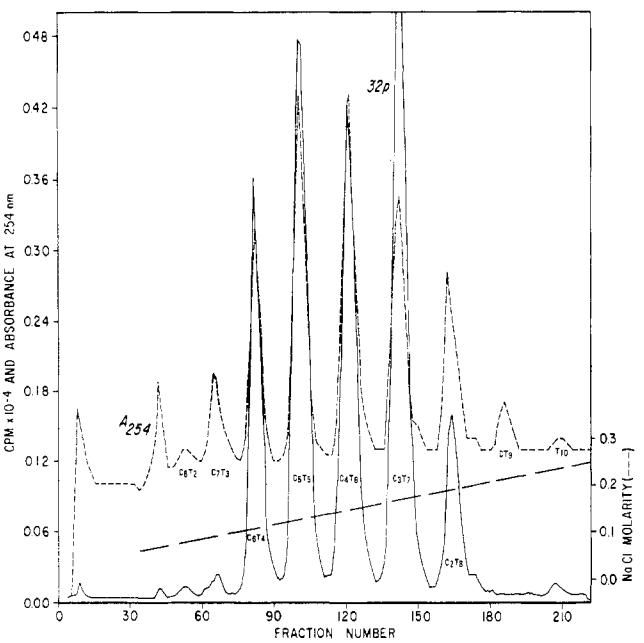

FIGURE 5: Fractionation of isostich 10. Legend as in Figure 4. Recovery from column = 90%.

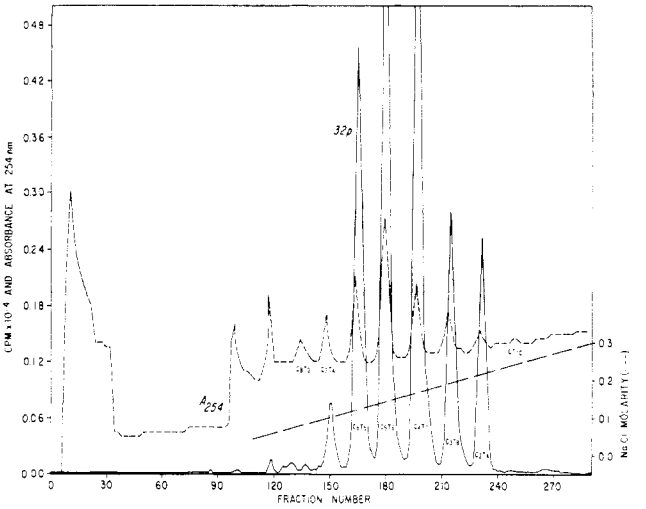

FIGURE 6: Fractionation of isostich 11. Legend as in Figure 4. Recovery from column = 92%.

corresponds to 1.37 tracts, hence they are both assigned the value of 1 tract/TP-84 genome.

The base composition of the components were determined in order to establish the correct C and T ratios. In six of the components digested a substance was found which cochromatographed on the nucleoside analyzer with riboadenosine. This material, which has a much smaller retention volume than 5-methyldeoxycytidine, was found in C_4T_6, C_3T_7, and C_5T_6 of the decanucleotide, C_7T_5 and C_5T_7 of the dodencanucleotide and C_5T_8 of the number 13. The ratio of this substance to T varied from 1:12 to 1:6. The highest proportion of the putative "rA" was found in component C_7T_5. The amount observed was sufficient to allow its occurrence in each dodecanucleotide chain at least once. Therefore, the component composition could be C_6T_5 rA. Since the amount of TP-84 DNA is low (100 μg) relative to calf thymus DNA (250 mg) in the original digest, it is quite certain that the rA-containing material derives from calf thymus DNA.

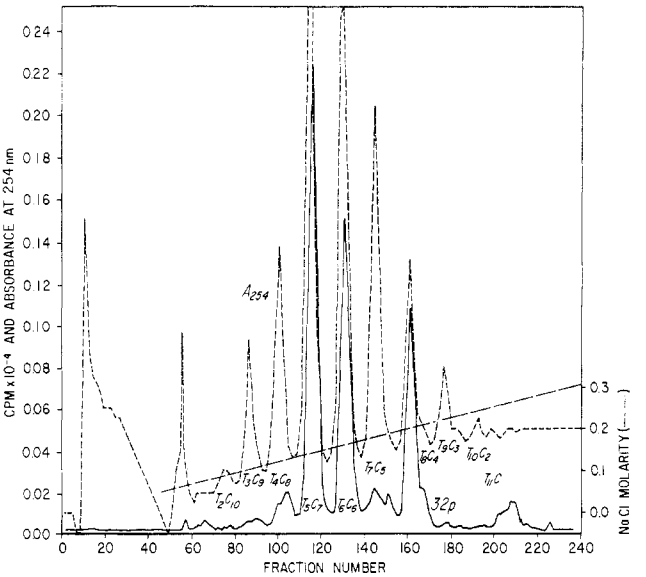

FIGURE 7: Fractionation of isostich 12. Legend as in Figure 4. Recovery from column = 98%.

TABLE II: Catalogs of Pyrimidine Tracts in TP-84 DNA.

Isostich	Component	Moles of Pyrimidine/100 g-atoms of DNA-P[a]	No. of Tracts/DNA Mol[b]	Isostich	Component	Moles of Pyrimidine/100 g-atoms of DNA-P[a]	No. of Tracts/DNA Mol[b]
1	C_{P2}	7.65 ± 0.13	7191	8	C_7T_{P9}	0.009	1
	T_{P2}	9.06 ± 0.13	8516		C_6T_{2p9}	0.021; 0.035	3
2	C_{2P3}	3.44 ± 0.13	1617		C_5T_{3p9}	0.042; 0.108	9
	CT_{P3}	7.59 ± 0.04	3567		C_4T_{4p9}	0.084; 0.106	11
	T_{2P3}	2.48 ± 0.16	1166		C_3T_{6p9}	0.273; 0.231	30
3	C_{3P4}	0.363 ± 0.027	114		C_2T_{6p9}	0.235; 0.202	26
	C_2T_{P4}	2.26 ± 0.06	708		CT_{7p9}	0.115; 0.097	12
	CT_{2P4}	3.91 ± 0.04	1225	9	C_8T_{p10}	0.003	
	T_{3P4}	1.28 ± 0.03	401		C_7T_{2p10}	0.024; 0.010	2
4	C_{4P5}	0.088 ± 0.029	21		C_6T_{3p10}	0.005; 0.024	2
	C_3T_{P5}	0.391 ± 0.022	92		C_5T_{4p10}	0.051; 0.054	6
	C_2T_{2P5}	1.35 ± 0.03	317		C_4T_{5p10}	0.112; 0.116	12
	CT_{3P5}	1.68 ± 0.02	395		C_3T_{6p10}	0.111; 0.111	12
	T_{4P5}	0.818 ± 0.021	192		C_2T_{7p10}	0.071; 0.076	8
5	C_{5P6}	0.036 ± 0.018	7		CT_{8p10}	0.045; 0.038	4
	C_4T_{P6}	0.124 ± 0.015	23		T_{9p10}	0.008	1
	C_3T_{2P6}	0.519 ± 0.018	98	10	C_8T_{2p11}	0.002 ±	
	C_2T_{3P6}	0.757 ± 0.008	142		C_7T_{3p11}	0.005 ± 0.002	
	CT_{4P6}	0.845 ± 0.016	159		C_6T_{4p11}	0.045 ± 0.002	4
	T_{5P6}	0.411 ± 0.016	77		C_5T_{5p11}	0.066 ± 0.002	6
6	C_{6P7}	0.004 ± 0.002	1		C_4T_{6p11}	0.063 ± 0.002	6
	C_5T_{P7}	0.036 ± 0.010	7		C_3T_{7p11}	0.077 ± 0.003	7
	C_4T_{2P7}	0.098 ± 0.010	15		C_2T_{8p11}	0.022 ± 0.002	2
	C_3T_{3P7}	0.398 ± 0.011	62	11	C_7T_{4p12}	0.006 ± 0.022	1
	C_3T_{4P7}	0.657 ± 0.010	103		C_6T_{5p12}	0.024 ± 0.002	2
	CT_{5P7}	0.473 ± 0.006	74		C_5T_{6p12}	0.033 ± 0.002	3
	T_{6P7}	0.179 ± 0.002	28		C_4T_{7p12}	0.033 ± 0.002	3
7	C_6T_{P8}	0.012 ± 0.002	2		C_3T_{8p12}	0.017 ± 0.002	1
	C_5T_{2P8}	0.044 ± 0.003	6		C_2T_{9p12}	0.014 ± 0.002	1
	C_4T_{3P8}	0.121 ± 0.007	16	12	C_8T_{4p13}	0.003; 0.004	
	C_3T_{4P8}	0.380 ± 0.003	51		C_7T_{5p13}	0.026; 0.029	2
	C_2T_{5P8}	0.456 ± 0.011	61		C_6T_{6p13}	0.018; 0.019	1
	CT_{6P8}	0.250 ± 0.005	34		C_5T_{7p13}	0.005; 0.003	
	T_{7P8}	0.076 ± 0.004	10		C_4T_{8p13}	0.014; 0.010	1
				13	C_7T_{6p14}	0.009	1
					C_6T_{7p14}	0.019	1
					C_5T_{8p14}	0.017	1
					C_4T_{9p14}	0.018	1
					C_3T_{10p14}	0.009	1

[a] Mean values for three separate analyses ±SEM on numbers 4, 5, 6, 10, and 11 and four separate analyses ±SEM on numbers 1, 2, and 3. [b] Number of tracts calculated using TP-84 DNA of mol wt 32 × 10^6; 47,000 base pairs.

Discussion

Bacteriophage TP-84 is specific for certain strains of *B. stearothermophilus*. TP-84 DNA has a molecular weight of 32 × 10^6 as determined by both sedimentation velocity (G. F. Saunders, unpublished data) and electron microscopy studies (A. Kleinschmidt and D. Lang, personal communication). Its nucleotide base composition is 42% G + C as determined by thermal denaturation, buoyant density measurements, and chemical analyses (Saunders and Campbell, 1965, 1966). No traces of minor bases were found in previous studies or the work presented here (Saunders and Campbell, 1965). Denatured DNA forms two bands in neutral CsCl density gradients, a pyrimidine-rich band and a purine-rich band (Saunders and Campbell, 1965). *In vivo* phage-induced RNA saturates 74% of the heavy or pyrimidine-rich DNA strand in DNA·RNA hybridization experiments (Shaw, 1969).

This distribution of A·T and G·C base pairs along the bacteriophage TP-84 DNA molecule was investigated. Three distinct compositional regions were indicated by the differential T_m (Figure 1). That these were of differing G + C content was evident from the T_m of each peak. The Felsenfeld denaturation graph (Figure 2) showed at least one region comprising 15% of the TP-84 genome which appeared to be rich in dG·dC pairs. Cytidylate stretches greater than six nucleotides long were not found by the tract analyses. Fewer dC-rich than

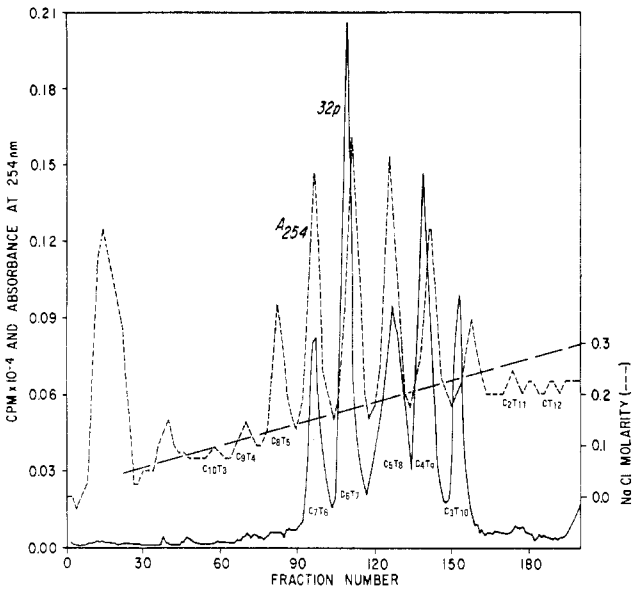

FIGURE 8: Fractionation of isostich 13. Legend as in Figure 4. Recovery from column = 93%.

dT-rich tracts were found in all isostichs. The distribution of total isostichs in TP-84 DNA correlate well with the pyrimidine catalogues of T7 DNA (Mushynski and Spencer, 1970a), λ DNA (Mushynski and Spencer, 1970b), and S13 RF DNA (Cerny *et al.*, 1969).

A total of 68 nonisomeric components were found in TP-84 DNA with no cytidylate runs longer than six nucleotides and no thymidylate runs longer than nine nucleotides. T_9 is the longest thymidylate run to be reported so far and T_8 is absent. The C_6 and T_9 moieties were found in amounts corresponding to one chain per TP-84 DNA molecule each. λ DNA and TP-84 DNA have 5 chains/genome each of isostichs 12 and 13, while T7 DNA has 7 chains in isostich 13.

The binding of poly(G) and poly(I,G) to denatured DNA was initially reported by Kubinski *et al.* (1966). Summers and Szybalski (1968a,b) postulated poly(G) binding sites found primarily on the *R* strand in phage T7 DNA to be deoxycytidylate-rich clusters 15–40 nucleotides long. The absence of deoxycytidylate runs in TP-84 longer than a hexanucleotide makes it difficult to explain the poly(I,G) binding sites known to occur in this DNA (Saunders *et al.*, 1969). Summers and Szybalski (1968a,b) suggest that these binding sites could be formed by a clustering of cytosine oligonucleotides three to five residues in length interrupted by purines or thymidine.

This explanation for poly(G) binding seems unlikely in view of the data presented by Mushynski and Spencer (1970a) on the distribution of pyrimidine oligonucleotides in the separated strands of T7 DNA. They showed an equal distribution of cytidylate tracts in isostichs three, four, and five in both the *r* and *l* strands of T7 DNA. However, they found isostichs 11, 12, and 13 almost exclusively on the *r* strand, the strand of T7 which binds poly(G) (Summers and Szybalski, 1968a) and poly(I,G) (Kubinski *et al.*, 1966). In λ DNA (Mushynski and Spencer, 1970b) the short-chain cytosine tracts were also distributed equally between the two separated strands. The slight bias found in isostichs 11, 12, and 13 was not sufficient to account for the extent of poly(G) binding of the *r* strand in this DNA.

Acknowledgment

We are grateful to Dr. J. T. Tidwell for his generous assistance in carrying out the nucleoside analyses.

References

Burton, K., and Peterson, G. B. (1960), *Biochem. J. 75*, 17.
Canellakis, E. S., and Mantsavinos, R. (1958), *Biochim. Biophys. Acta 27*, 643.
Cerny, R., Cerna, E., and Spencer, J. H. (1969), *J. Mol. Biol. 46*, 145.
Clausen, T. (1968), *Anal. Biochem. 22*, 70.
Darby, G., Dumas, L. B., and Sinsheimer, R. L. (1970), *J. Mol. Biol. 52*, 227.
Felsenfeld, G., and Hirschman, S. Z. (1965), *J. Mol. Biol. 13*, 407.
Felsenfeld, G., and Sandeen, G. (1962), *J. Mol. Biol. 5*, 587.
Gilbert, W., and Müller-Hill, B. (1966), *Proc. Nat. Acad. Sci. U. S. 56*, 1891.
Hall, J. B., and Sinsheimer, R. L. (1963), *J. Mol. Biol. 6*, 115.
Hirschman, S. Z., Felsenfeld, G. (1966), *J. Mol. Biol. 16*, 347.
Hirschman, S. Z., Gellert, M., Falkow, S., Felsenfeld, G. (1967), *J. Mol. Biol. 28*, 464.
Kubinski, H., Opara-Kubinska, Z., Szybalski, W. (1966), *J. Mol. Biol. 20*, 313.
Mushynski, W. E., Spencer, J. H. (1970a), *J Mol. Biol. 52*, 91.
Mushynski, W. E., and Spencer, J. H. (1970b), *J. Mol. Biol. 52*, 107.
Petersen, G. B., and Reeves, J. M. (1966), *Biochim. Biophys. Acta 129*, 438.
Ptashne, M. (1967), *Proc. Nat. Acad. Sci. U. S. 57*, 306.
Rudner, R., Shapiro, H. S., and Chargaff, E. (1966), *Biochim. Biophys. Acta 129*, 85.
Rushizky, G. W., and Sober, H. A. (1964), *Biochem. Biophys. Res. Commun. 14*, 276.
Saunders, G. F., and Campbell, L. L. (1965), *Biochemistry 4*, 2836.
Saunders, G. F., and Campbell, L. L. (1966), *J. Bacteriol. 91*, 340.
Saunders, G. F., Reese, W. N., Bergendahl, J., Mandel, M. (1969), *Biochemistry 8*, 1222.
Shapiro, H. S., and Chargaff, E. (1964), *Biochim. Biophys. Acta 91*, 262.
Shapiro, H. S., Rudner, R., Miura, K., and Chargaff, E. (1965), *Nature (London) 205*, 1068.
Shaw, M. Van C. (1969), Ph.D. Thesis, University of Illinois.
Spencer, J. H., Cape, R. E., Marks, A., and Mushynski, W. E. (1968), *Can. J. Biochem. 46*, 627.
Spencer, J. H., and Chargaff, E. (1963), *Biochim. Biophys. Acta 68*, 18.
Steinberg, R. A., Ptashne, M. (1971), *Nature (London) 230*, 76.
Summers, W. C., Szybalski, W. (1968b), *Biochim. Biophys. Acta 166*, 371.
Summers, W. C., Szybalski, W. (1968a), *Virology 34*, 9.
Szybalski, W., Kubinski, H., and Sheldrick, P. (1968), *Symp. Quant. Biol. 31*, 123.
Uziel, M., Koh, C. K., and Cohn, W. E. (1968), *Anal. Biochem. 25*, 77.