Dabelsteen, E., Vedtofte, P., Hakomori, S., & Young, W. W. (1982) *J. Invest. Dermatol. 79*, 3–7.
Folch-Pi, J., Arsov, S., & Meath, J. A. (1951) *J. Biol. Chem. 191*, 819–831.
Fukuda, M. N., Fukuda, M., & Hakomori, S. (1979) *J. Biol. Chem. 254*, 5458–5465.
Fukushi, Y., Hakomori, S., Nudelman, E., & Cochran, N. (1984) *J. Biol. Chem. 259*, 4681–4685.
Hakomori, S. (1981) *Semin. Hematol. 18*, 39–62.
Hakomori, S., Stellner, K., & Watanabe, K. (1972) *Biochem. Biophys. Res. Commun. 49*, 1061–1068.
Kabat, E. A. (1982) *Am. J. Clin. Pathol. 78*, 281–292.
Kannagi, R., Nudelman, E., Levery, S. B., & Hakomori, S. (1982) *J. Biol. Chem. 257*, 14865–14874.
Kannagi, R., Stroup, R., Cochran, N. A., Urdal, D. L., Young, W. W., Jr., & Hakomori, S. (1983a) *Cancer Res. 43*, 4997–5005.
Kannagi, R. Levery, S. B., Ishigami, F., Hakomori, S., Shevinsky, L. H., Knowles, B. B., & Solter, D. (1983b) *J. Biol. Chem. 258*, 8934–8942.
Karlsson, K.-A. (1976) in *Glycolipid Methodology* (Witting, L. A., Ed.) pp 97–122, American Oil Chemists' Society Press, Champaign, IL.
Kisailus, E. C., & Kabat, E. A. (1978) *J. Exp. Med. 147*, 830–843.
Köhler, G., & Milstein, C. (1975) *Nature (London) 256*, 495–497.
Lloyd, K. O., Kabat, E. A., & Layug, F. (1966) *Biochemistry 5*, 435–445.
Magnani, J. L., Smith, D. F., & Ginsburg, V. (1980) *Anal. Biochem. 109*, 399–402.
McKibbin, J. M. (1976) in *Glycolipid Methodology* (Witting, L. A., Ed.) pp 77–95, American Oil Chemists' Society Press, Champaign, IL.
McKibbin, J. M. (1978) *J. Lipid Res. 19*, 131–147.
McKibbin, J. M., Spencer, W. A., Smith, E. L., Mansson, J.-E., Karlsson, K.-A., Samuelsson, B. E., Li, Y.-T., & Li, S.-C. (1982) *J. Biol. Chem. 257*, 755–760.
Moreno, C., Lundblad, A., & Kabat, E. A. (1971) *J. Exp. Med. 134*, 439–457.
Siddiqui, B., & Hakomori, S. (1971) *J. Biol. Chem. 246*, 5766–5769.
Slomiany, A., & Slomiany, B. L. (1975) *Biochim. Biophys. Acta 388*, 135–145.
Smith, E. L., McKibbin, J. M., Breimer, M. E., Karlsson, K.-A., Pascher, I., & Samuelsson, B. E. (1975) *Biochim. Biophys. Acta 398*, 84–91.
Vance, W. R., Shook, C. P., III, & McKibbin, J. M. (1966) *Biochemistry 5*, 435–445.
Watkins, W. M. (1980) *Adv. Hum. Genet. 10*, 1–136.
Young, W. W., Jr., MacDonald, E. M. S., Nowinski, R. C., & Hakomori, S. (1979) *J. Exp. Med. 150*, 1008–1019.
Yu, R. K., & Ledeen, R. W. (1972) *J. Lipid Res. 13*, 680–686.

Analysis of DNA Sequences Using a Single Chemical Cleavage Procedure

Barbara J. B. Ambrose and Reynaldo C. Pless*
Department of Chemistry, The Ohio State University, Columbus, Ohio 43210
Received December 10, 1984

ABSTRACT: A novel approach to sequence analysis of end-labeled, defined DNA fragments, using a single chemical cleavage procedure and electrophoretic separation in a single lane, has been developed. Prolonged treatment with hot aqueous piperidine results in partial cleavage of the DNA at all positions; the relative propensity for this cleavage is different for the various bases in the DNA. The hydrolysate is resolved on a DNA sequencing gel, and the distribution of radioactivity in the electrophoretic lane is analyzed (a) in terms of differential peak heights of the radioactive bands and (b) in terms of the spacings between successive bands. Simultaneous application of these two base-characteristic criteria allows the deduction of the nucleotide sequence with an accuracy approaching that of the established four-lane methods of DNA sequencing.

In the chemical cleavage method for DNA sequencing as developed by Maxam & Gilbert (1977, 1980), the DNA fragment of interest is provided with a radioactive tag specifically at one terminus and then subjected, in parallel, to *four* separate DNA cleavage procedures, which differ in their base specificities. Parallel resolution of the four reaction mixtures by polyacrylamide gel electrophoresis followed by visualization of the radioactive bands produces a four-lane pattern from which the nucleotide sequence of the fragment can be read off directly.

A considerable simplification would be achieved if the four parallel cleavage procedures were to be substituted by *one* cleavage protocol which should (a) be capable of severing the DNA backbone at *each* nucleotide position and (b) have distinct and characteristic propensities for DNA cleavage at the four different bases, adenine, guanine, cytosine, and thymine (A, G, C, and T, respectively).[1] The DNA sequence would then be deduced from the succession of radioactive bands of varying intensities in *one* electrophoretic lane. Such a method would also have good potential for automation of DNA sequencing.

The present is a report on a procedure which uses such differential cleavage at the various DNA bases, in conjunction

[1] Abbreviations: A, G, C, and T, adenine sites, guanine sites, cytosine sites, and thymine sites in DNA, respectively; ATP, adenosine 5′-triphosphate; EDTA, ethylenediaminetetraacetic acid; Tris, tris(hydroxymethyl)aminomethane; DEAE, diethylaminoethyl.

0006-2960/85/0424-6194$01.50/0

with nucleotide-specific decrements in electrophoretic mobility between successive members of a nested set of polynucleotides, as the criteria to determine nucleotide sequence by using a single electrophoretic lane.

Materials and Methods

Materials. Chemicals were obtained from the following sources: Tris(hydroxymethyl)aminomethane (TRIZMA base) and ammonium persulfate, Sigma; acrylamide, *N,N'*-methylenebis(acrylamide), and *N,N,N',N'*-tetramethylethylenediamine, Bio-Rad Laboratories; urea, piperidine, and formamide, Fisher Scientific Co.; sodium borate, Mallinckrodt Chemical Works; formic acid, Ashland Chemical Co.; sodium chloride, MCB Reagents; [γ-^{32}P]ATP (6000 Ci/mmol), New England Nuclear.

Purified pBR322 DNA fragments containing a specifically labeled 5′ terminus were the gift of Dr. Lee F. Johnson and Sylvia Perryman. These had been prepared by established procedures (Maxam & Gilbert, 1980) by restriction of full-length pBR322 DNA with a specific endonuclease, 5′-dephosphorylation with calf intestinal phosphatase, 5′-end labeling by rephosphorylation with [γ-^{32}P]ATP using T4 polynucleotide kinase, digestion with a second restriction endonuclease, resolution of the fragments by electrophoresis through a 6% polyacrylamide gel, and elution of desired fragments (double stranded, but bearing a single, defined 5′-terminal label).

Methods. Hydrolyses in 0.5 M aqueous piperidine containing varying concentrations of NaCl and unlabeled sonicated salmon sperm DNA at 0.2 mg/mL were performed in 16-μL volumes in unsiliconized, sealed glass capillaries heated at 90 °C in a silicone oil bath. After fracture of the tubes, each sample was lyophilized, twice again lyophilized from 10 μL of water, and taken up in a small volume of a loading buffer containing aqueous formamide, tracking dyes, sodium hydroxide, and EDTA (Maxam & Gilbert, 1980). This was heated for 2 min at 90 °C, and a 3-μL sample (containing approximately 10000 cpm of radioactivity) was placed in the loading well. The vertical electrophoresis gels (40 cm long, 0.2 mm thick) were made from 19% acrylamide, 1% *N,N'*-methylenebis(acrylamide), and 7 M urea in running buffer (100 mM Tris–borate, pH 8.3, and 2 mM EDTA). Gels were run at an electrical field of approximately 60 V/cm; the surface temperature of the gel mold was approximately 40 °C. After opening of the mold, the gel was covered on one side with saran wrap, and Kodak XAR-5 film was exposed to it, at −70 °C, in the presence of an intensifying screen (Cronex Lightning Plus, Du Pont). After development, the film was scanned on a Zeineh Soft Laser scanning densitometer interfaced with an Apple IIe computer and the densitometric profile plotted by a printer.

Results

Upon prolonged treatment with 0.5 M aqueous piperidine at 90 °C, DNA is cleaved at *all* positions, i.e., at A, G, C, and T sites, with comparable frequency (Pless & Bessman, 1983). After treatment of a 5′-end labeled, defined DNA fragment in this manner, resolution of the hydrolysate by polyacrylamide gel electrophoresis and autoradiography will produce a ladder pattern displaying the *entire* set of possible polynucleotides containing the original 5′ terminus. The relative intensity of the different bands will reflect the propensity of the DNA to cleavage at the various positions.

To assess the potential of this reaction for a chemical cleavage procedure for DNA sequencing requiring a single sample treatment and a single electrophoretic lane, we have examined several variants of this reaction, using as test examples defined fragments of plasmid pBR322, bearing 5′-terminal ^{32}P labels.

Our results show that upon prolonged treatment with 0.5 M piperidine at 90 °C, DNA is most susceptible to cleavage at A sites, G sites, and C sites (the actual preference depending on salt concentration) and least prone to cleavage at T sites. Apart from the use of base-characteristic propensities for cleavage (i.e., band intensities in the gel electrophoretogram), our method of interpretation is based upon an additional criterion which we name *rhythm*. By this term, we mean the alternation of long and short separations relating successive polynucleotide bands, which is nucleotide specific; as observed by Maxam & Gilbert (1977), in sequencing ladders created by chemical cleavage of DNA, bands denoting G or T show a relatively large separation from the next larger polynucleotide (the next higher step in the ladder), while for the bands denoting A or C this separation is relatively small.

This nucleotide-specific alternation in band spacing is obviously superimposed on the systematic variation in band separation with polynucleotide length which is observed in polyacrylamide gel electrophoresis (Sealey & Southern, 1982). This does not represent an insuperable complication: both in visual inspection of autoradiograms and in a more quantitative treatment (as shown below), the approximately inverse dependence of mobility on chain length (Southern, 1979) can be taken into account, and a particular step length can be characterized as short or long relative to the average step length in the vicinity of the band of interest.

Figure 1 shows the autoradiographed gel electrophoretograms of hydrolysates obtained by heating a 5′-end-labeled 207-mer [positions 379–173 in the sequence of Sutcliffe (1979)] in 0.5 M aqueous piperidine at 90 °C for 5 h, in the presence of different concentrations of added NaCl. Four such hydrolyses were run in parallel, with NaCl concentrations of 0.10, 0.30, 1.0, and 2.5 M. Three sets of electrophoretograms are shown in Figure 1, one set (lanes A1, A2, A3, and A4) in which the hydrolysates were electrophoresed for a relatively short time, so that all oligonucleotides starting from a chain length of L = 2 phosphates are seen on the gel, and two other sets in which the hydrolysates were run for longer times, so that all polynucleotides with a chain length shorter than L = 41 phosphates were electrophoresed off the gel in C1, C2, C3, and C4, and all polynucleotides with a chain length shorter than L = 50 phosphates were run off in D1, D2, D3, and D4.

On examination of lane C2 (hydrolysis in the presence of 0.3 M NaCl), one sees that the 2nd, 5th, 9th, and 15th bands, counting from the bottom, are of very weak intensity. These weak bands denote T positions, according to the sequence of Sutcliffe (1979). The 1st, 3rd, 12th, 13th, 16th, 20th, and 21st bands are strong; these are A bands. For these bands, the separation from the next higher band in the ladder is smaller than for the T bands. The band spacings seen for T bands and for A bands are taken as a standard to which other band spacings in the adjoining regions of the gel can be compared: while T spacings qualify as moderately long, A spacings qualify as moderately short for a given region of the gel. The rest of the bands in this series are of intermediate intensity and represent G and C sites; these are assigned on the basis of rhythm: for C bands (8th, 10th, 11th, and 18th bands), the spacing is even shorter than for A bands, while for G bands (4th, 6th, 7th, 14th, 17th, and 19th bands), the spacing is longer than for T bands. Thus, for this series of bands, the simultaneous application of the two criteria of relative band intensity and relative band spacing allows the correct assign-

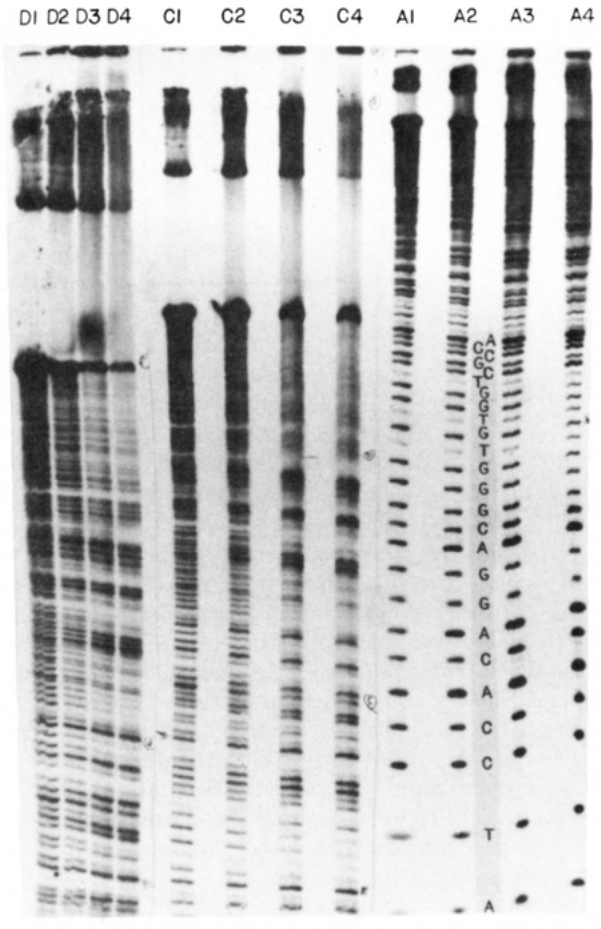

FIGURE 1: Autoradiogram of a gel electrophoretogram of hydrolysates obtained from the 5′-end-labeled 207-mer (positions 379–173) by treatment with 0.5 M aqueous piperidine at 90 °C for 5 h. Electrophoresis times were 4 h for the lanes denoted A, 16 h for the lanes denoted C, and 20 h for the lanes denoted D. Concentrations of NaCl during the aqueous piperidine treatment were 0.1, 0.3, 1.0, and 2.5 M for the lanes denoted 1, 2, 3, and 4, respectively. The loading wells are at the top of the picture; electrophoretic movement was downward. For the bottommost 25 bands in lane A2, the band identification is given in the figure.

ment of the bases: A, strong band, relatively short separation to the next longer polynucleotide; G, medium band, relatively large separation; C, medium band, relatively short separation; T, weak band, relatively large separation.

Using this prescription, we have read the sequence up to a position 120 nucleotides removed from the labeled end, in lane D2 (Figure 1). A detail of this electrophoretic lane, showing the bands corresponding to the nucleotides between positions 311 and 293 in Sutcliffe's sequence, is shown in Figure 2.

For sequences very close to the labeled terminus, interpretation of the band pattern is complicated by three effects. First, the band shape is considerably distorted, probably due to comigration of salt in that region of the electrophoretogram (compare to the bottom portions of lanes A3 and A4, which contained more salt). Second, the first three or four bands often show abnormal intensities (e.g., the second band from the bottom in lane A2, which is much stronger than expected for a T band), possibly due to end effects in the hydrolysis reaction. Third, in contrast to the slow change in average band spacing observed in the regions containing the longer oligonucleotides (e.g., the bottom portion of lane C2), the systematic change in spacing becomes drastic for the very short oligomers. Nonetheless, except for the first two bands (A and T at the bottom of lane A2), spacing can be correctly interpreted, as follows. In absolute terms, the spacing of the third band in

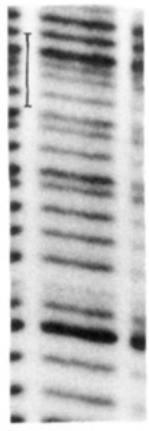

FIGURE 2: Detail of lane D2 of Figure 1, showing the bands corresponding to positions 311–293, according to Sutcliffe (1979). Reading upward, the sequence is GGACTGGGCGGCGGCCAAA. The bar indicates an actual length of 1 cm.

lane A2 is only slightly larger than the spacing of the fifth band (which is an A band, as seen from its intensity); since for the very short oligonucleotides a decrease in chain length by even one nucleotide unit results in a large increase in the average separation, the spacing of the third band qualifies as shorter than an A spacing. By similar reasoning, the spacing of the fourth band, which in absolute terms is identical with the spacing of the fifth band, qualifies as shorter than an A band. The sixth band obviously has a shorter spacing than do local A bands. The spacing of the eighth band is identical with the spacing of the preceding seventh band (an A band, by intensity); thus, it is interpreted as relatively larger than an A band. The 9th band has an electrophoretic interval which is far larger (by 40%) than the spacing of the following A band (the 10th band); therefore, the spacing for the 9th band is interpreted as relatively long.

A comparison of the hydrolytic behavior of the DNA in the presence of different concentrations of NaCl, in terms of the overall rate of degradation, shows that the hydrolysis becomes more rapid with increasing salt concentration. This is apparent from the gradual decrease in the intensity of the parent peak in going from 0.1 M to 2.5 M salt (compare lanes C1, C2, C3, and C4). Furthermore, in the high-salt hydrolysates, the large fragments show very low intensities (in the region below the parent peaks in lanes C3 and C4) while the short fragments are very pronounced (at the bottom of lanes A3 and A4; this increase of the smaller fragments at the expense of the larger fragments shows that under these salt conditions multiple cleavage occurred to a significant extent during the hydrolysis.

When the effect of varying salt concentrations on the interpretability of the sequence on the basis of different peak heights is compared, 0.3 and 1.0 M NaCl are most favorable in ensuring that A bands will clearly be more intense than C bands (a critical requirement since A bands and C bands cannot reliably be distinguished by spacing) and that G bands are sufficiently prominent over T bands (again critical since G bands and T bands have similar spacing). This is best seen from a comparison of the densitometric profiles in panels A1, A2, A3, and A4 in Figure 3, which correspond to the electrophoretic lanes A1, A2, A3, and A4 in Figure 1. Disregarding the first four bands (corresponding to the shortest oligonucleotides), one sees that in the hydrolysate obtained at 0.3 M NaCl (panel A2) A bands are distinctly stronger than C bands, which are equal in intensity to G bands; T bands are clearly the weakest. The hydrolysate obtained with 1.0 M NaCl (panel A3) shows similar relationships in the peak heights, except that C bands are here somewhat stronger than

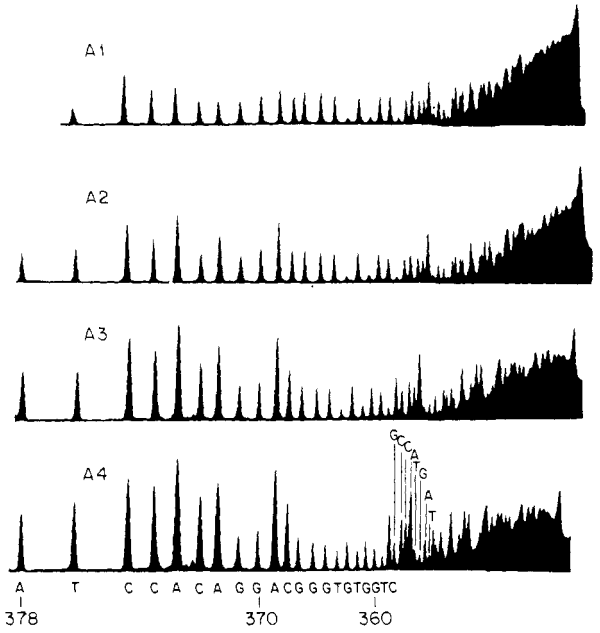

FIGURE 3: Densitometric profiles of lanes A1, A2, A3, and A4 of Figure 1. The concentrations of NaCl during the hydrolysis of hot aqueous piperidine were as follows: A1, 0.10 M; A2, 0.30 M; A3, 1.0 M; A4, 2.5 M. In each case, the bottom of the gel is at the left, and the major band at the right end is the parent peak. Nucleotide sequence and position numbers according to Sutcliffe (1979) are indicated.

G bands. When hydrolysis is performed in the presence of 2.5 M NaCl (panel A4), G bands are relatively reduced to a point where their reliable distinction from T bands is no longer possible. In hydrolysis at 0.1 M NaCl (panel A1), DNA has similar propensities for cleavage at A sites, G sites, and C sites, so that the critical distinction of A bands from C bands on the basis of different peak heights becomes unreliable. A similar situation obtains at the much lower salt concentration of 0.01 M (data not shown).

For the hydrolysate obtained with 0.3 M NaCl, Figure 4 shows, in panels B2, C2, and D2, the densitometric profiles measured after electrophoreses of longer durations. Panel B2 shows polynucleotides starting from a chain length of $L = 17$ phosphates (original autoradiogram not shown). Panel C2 shows polynucleotides starting from a chain length of $L = 41$ phosphates (this corresponds to lane C2 in Figure 1), and panel D2 shows polynucleotides starting from a chain length of $L = 50$ phosphates (this corresponds to lane D2 in Figure 1.)

Between positions 375 and 260 (i.e., between positions 5 and 120 nucleotides removed from the labeled 5′ end), our reading agrees with the sequence given by Sutcliffe (1979), except in two locations. First, the peak at position 351 (A in Sutcliffe's sequence) has a band spacing characteristic of A; in peak height, however, it appears as a G band or a C band (see Figure 4, panel B2); this reduced susceptibility of A to cleavage must be due to methylation since the sequence GATC is methylated at A in the *Escherichia coli* host. (Upon piperidine treatment in 0.1 M NaCl, this methyladenine band is slightly stronger than T bands; at 1.0 or 2.5 M NaCl, the methyladenine band is intermediate in intensity between G bands and C bands.) Second, for C 317 and G 316 in Sutcliffe's sequence, we see a single peak, with an intensity characteristic of an A band (Figure 4, panels C2 and D2). We have confirmed (data not shown) by analysis according to Maxam & Gilbert (1980) that the DNA fragment used by us agrees completely with the sequence of Sutcliffe, at least from position 373 to position 286. The discrepancy at C 317 and G 316

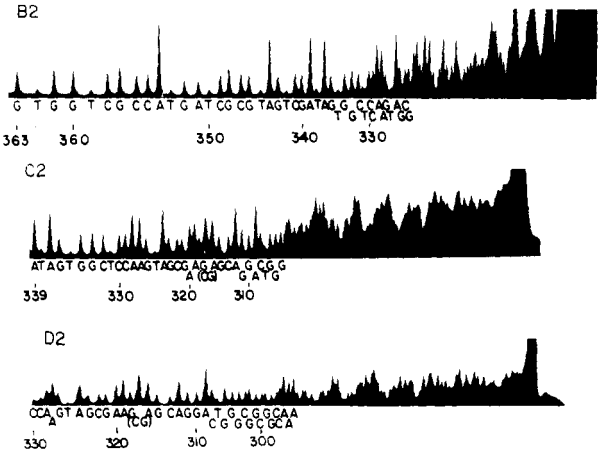

FIGURE 4: Densitometric profiles for the hydrolysate of the end-labeled 207-mer with 0.5 M aqueous piperidine and 0.3 M NaCl. Electrophoresis times were 8 (B2), 16 (C2), and 20 h (D2). Nucleotide sequence and position numbers according to Sutcliffe (1979) are indicated for each panel.

arises from very close overlap of these two bands (the comigration of the bands for C 317 and G 316 can also be seen in our analysis according to Maxam and Gilbert). Electrophoresis at higher temperature (55 °C, measured at the glass surface) shows these positions as one slightly broadened band, indicating incipient separation (data not shown). This is a case of band compression due to loop formation (Maxam & Gilbert, 1980): for the local sequence GCGA (positions 318–315), two complementary tetranucleotide sequences are available on the 5′ side, viz., TCGC (positions 359–356) and TCGC (350–347). In line with this interpretation is the observation that several subsequent bands (A 315, G 314, C 313, A 312, and G 311) show uncommonly large spacings: this is usually observed after regions of band compression caused by loop formation (Maxam & Gilbert, 1980). It is interesting to note that the almost immediately preceding identical tetranucleotide sequence GCGA (323–320) is not similarly affected by band compression: this indicates that loop formation is only stable enough to be significant when *both* GCGA sequences interact with *both* TCGC sequences.

Figure 5 shows sequencing lanes obtained from several different 5′-end-labeled fragments of pBR322 DNA, after treatment with 0.5 M aqueous piperidine and 0.3 M NaCl at 90 °C for 5 h. Figure 5 also shows, in lane A, a hydrolysate obtained by treatment with 0.5 M aqueous piperidine and 0.5 M NaCl at 110 °C for 1.5 h. While all other electrophoretic separations described in this report used 20% polyacrylamide gels, the electrophoretic separations shown in Figure 5, lanes F and G, were performed in 12% and 8% polyacrylamide gels, respectively. In all these instances shown in Figure 5, extensive sequences were correctly read by applying the rules pertaining to band intensities and spacings which were derived above, with the exception of the methyladenine sites in GATC sequences (A 3737 and A 351).

An approximately linear relationship exists between polynucleotide chain lengths and their inverse mobilities on DNA sequencing gels (Southern, 1979). Therefore, a calculation of the increments of the function $1/m$ (i.e., the inverse of the distance migrated, measured from the electrophoretic origin) for successive bands should allow a more quantitative assessment of the *nucleotide-specific* component in the band spacings, divorced from the systematic decrease in band separations with increasing chain length. Figure 6 shows the plot of this quantity $[\Delta(1/m)]$ vs. polynucleotide length (L, equal to the number of phosphates in the polymer) for lane A2 in

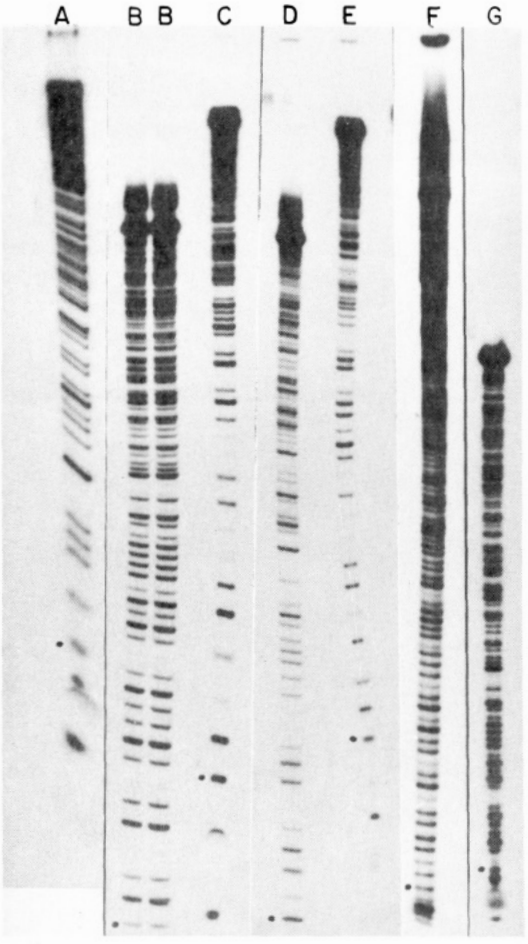

FIGURE 5: Autoradiograms of five separate gel electrophoretograms of hydrolysates obtained from various 5′-end-labeled fragments of pBR322 DNA [lane A, fragment 653–380, according to Sutcliffe (1979); lanes B (identical), C, and G, 3867–3731; lanes D and E, 3728–3864; lane F, 379–173. Hydrolysis was in 0.5 M aqueous piperidine and 0.3 M NaCl at 90 °C for 5 h except for lane A, where it was in 0.5 M aqueous piperidine and 0.5 M NaCl at 110 °C for 1.5 h. Polyacrylamide gel concentration was 20%, except for lanes F (12%) and G (8%). The gels were run at a temperature of approximately 40 °C, except for lane A which was run at approximately 50 °C. Reading up from the bands marked by dots, the sequences are as follows: lane A, CTCTCCCTTATGCGACTCCTGCATTAGGAA...; lanes B, GAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA...; lane C, AATGACTTGGTTGAGTACTCACCAGT...; lane D, AGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTA...; lane E, CCGATCGTTGTCAGAAGTAAGTTGGCCGCA...; lane F, GATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA...; lane G, CTCACCAGTCACAGAAAAGCATCTTACGGATGGCA....

Figure 1. The $\Delta(1/m)$ values for the first, very short fragments (the nucleoside diphosphate denoting A 378 and the dinucleoside triphosphate denoting T 377) are exceptionally high. For higher values of L, the values of $\Delta(1/m)$ for any one of the four canonical nucleotides are seen to vary within fairly narrow boundaries; they are, therefore, a useful criterion for the identification of the bands in the electrophoretic ladder. A distinct deviation of some points in this graph to lower $\Delta(1/m)$ values may be caused by formation of weak hairpin structures: thus, for G at $L = 10$ (i.e., G 370), the reason could be binding of CC(376–375) to GG(371–370); for A at $L = 11$ (i.e., A 369), binding of TCC(377–375) to GGA(371–369); for G at $L = 20$ (i.e., G 360), binding of CCACA(376–372) to TGTGG(364–360). Even with these deviations, values of $\Delta(1/m)$ for G bands and C bands (which cannot be distinguished by peak heights after hydrolysis in the presence of 0.3 M NaCl) are characteristically different over the whole range

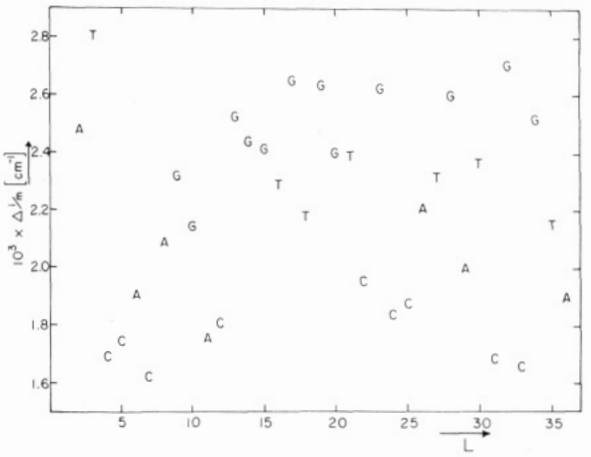

FIGURE 6: Plot of the increment in inverse mobility in going from the polynucleotide of chain length L to the polynucleotide of chain length $L + 1$ vs. chain length L. The coordinate points are marked by letter symbols identifying in each case the particular base signalized by the radioactive polynucleotide of chain length L.

of L values from $L = 4$ to $L = 36$.

A similar situation obtains in a plot (not shown) of $\Delta(1/m)$ vs. L for the 17-mer through the 70-mer for the densitometric scan shown in panel B2 in Figure 4, except that there is no abnormality in the $\Delta(1/m)$ values at the bottom of the gel and that there is the expected dramatic increase in $\Delta(1/m)$ values for several nucleotide units downstream from G 316 (i.e., past the band compression region).

Discussion

The four-lane sequencing method of Maxam & Gilbert (1977, 1980) logically uses a qualitative criterion—the presence or absence of a radioactive band in a certain electrophoretic lane—to answer the qualitative question of the identity of a base in a given position. A method which requires a more quantitative analysis to yield the same qualitative information must obviously redeem itself by other advantages to be of interest. These are the following: (a) The cleavage procedures are simplified. The method described here entails the treatment of *only one sample* of the end-labeled DNA fragment with 0.5 M aqueous piperidine at 90 °C at 0.3 or 1.0 M NaCl, for 5 h. With a temperature of 110 °C and 0.5 M NaCl, a hydrolysis time of 1.5 h was sufficient to produce a satisfactory electrophoretic ladder (lane A, Figure 5). The procedure is very reproducible and rather insensitive to the exact conditions and materials used: moderate variations in the salt concentration and the use of piperidine of different grades of purity do not appreciably affect the results. (b) The procedure avoids the use of such toxic substances as dimethyl sulfate and hydrazine. (c) As only one lane is used, the requirement for DNA sample is reduced. (d) Analysis in one lane eliminates the problems of band ordering which arise in four-lane analysis due to the curvature of the electrophoretic front. The one-dimensional disposition of all the relevant information (peak height and spacing) makes this method particularly promising for automated analysis of the radioactivity distribution on the gel in terms of a DNA sequence; we are at present developing these methods.

The effect of impurity of the DNA sample on sequence analysis by the one-lane method should be considered. DNA contaminants smaller than the fragment under analysis will, in their undegraded form, seriously interfere with the reading of the ladder. Such DNA contaminants are, however, easily removed during electrophoretic purification of the DNA fragment for analysis, unless they are of closely similar mo-

bility; in this latter case, they can only interfere, in their undegraded form, with the reading of the region close to the parent band. The degradation products of such contaminants would, however, interfere along the entire length of the electrophoretic lane: as the ratio of intensities of A bands to T bands is about 10:1, a DNA contaminant present to the extent of 10% of the fragment under analysis would produce a contaminating ladder in which the A bands would be similar in intensity to the T bands of the ladder of interest. Therefore, the one-lane method requires a sample in which the DNA impurities amount to maximally a few percent.

Compared to the four-lane methods, the present approach is more dependent on good band resolution. This is due (a) to the fact that the segregation of bands into four separate lanes relieves crowding and (b) to the fact that in the procedure presented here the electrophoretic pattern is examined not only for a sequential ordering of bands but also for an asessment of band spacings as relatively short or relatively long. In general, this presents no problem, as the band separations characteristic for G, T, A, and C bands relate approximately as 1.3:1.2:1.1:1.0, placing a difference of 30% between the electrophoretic intervals for the G and C bands, which cannot be distinguished by intensity. Thus, even when an electrophoretic lane shows considerable imperfection, e.g., the strongly slanted bands in lane A in Figure 5, distinction of G and C bands by spacing is straightforward.

Interpretation of band spacing is difficult only in the region containing the very short oligonucleotides. How close to the labeled end the sequence can be read depends on the analyst's sense of rhythm; this develops with experience, and an examination of the electrophoretograms shown in this report will be helpful in this respect. Interpretation of the first three bands is usually senseless.

In comparison to the established methods, our approach is more seriously affected by abnormalities in the electrophoretic mobilities, e.g., band compression. In the sequences examined by us, which aggregate a total of approximately 800 nucleotides, this phenomenon appeared in a marked form only once: positions C 317 and G 316 in the sequence of Sutcliffe (1979), which appear as a single band in our analysis. The problem was clearly signaled by the extraordinarily large spacings seen for the immediately following bands. Obviously, with the present one-lane approach, the interpretation of several bands becomes uncertain in a region of this type, and the uncertainty can only be resolved safely by sequencing through the region in the opposite direction. This is also necessary if the method of Maxam & Gilbert (1980) is used on this fragment, as even with elevated electrophoresis temperatures the bands signaling C 317 and G 316 do not separate sufficiently to allow reliable ordering.

The use of mobility shifts in one-dimensional separation systems for the characterization of bases in a sequence was first examined by Brownlee & Sanger (1967) for the case of ribooligonucleotides resolved by DEAE paper ionophoresis at pH 3.5. For polydeoxyribonucleotides separated by polyacrylamide gel electrophoresis, Maxam & Gilbert (1977) pointed out the base specificity of the band spacings and gave a physical rationale for the observation. The present report shows that this base specificity holds over a large range of polynucleotide chain lengths and for polyacrylamide gels of different concentrations.

While more elaborate plots of electrophoretic mobilities [e.g., $1/(m - m_0)$ vs. chain length; see Southern (1979)] give a better linearity, we have chosen in our quantitative analysis of band spacings the simple inverse mobility, $1/m$. This form already sufficiently compensates for the systematic change in spacing with chain length to make the decrements in mobility [now expressed in terms of the differential parameter $\Delta(1/m)$] characteristic for the different bases over a large range of polynucleotide lengths (Figure 6).

The cleavage of the DNA backbone in hot aqueous piperidine does not evince sequence dependence; i.e., the propensity for cleavage is the same at all positions occupied by a certain base. This is expected in view of the random-coil conformation of the DNA under the temperature and pH conditions of this hydrolysis.

Our understanding of the chemistry underlying the differential cleavage of DNA by 0.5 M aqueous piperidine at 90 °C is deficient. We have found (data not shown) that bands denoting purines obtained by this procedure comigrate precisely with the corresponding bands obtained with the conventional (A + G) cleavage procedure (Maxam & Gilbert, 1980) run in a parallel channel. This indicates that the identical fragment results in both instances. Thus, in both procedures, the sequence d-p*XpYpZpPur..., where Pur stands for A or G, gives rise to the fragment d-p*XpYpZp. The sequence of events is probably identical with that occurring in the case of the various base-specific reactions described by Maxam & Gilbert (1980): removal of the heterocyclic base, followed by two successive β-eliminations in the deoxyribose moiety, releasing both the 3′ remainder and the 5′ remainder of the DNA fragment, with their 5′-terminal phosphate and their 3′-terminal phosphate, respectively, still attached. If this general scheme for the cleavage events holds, the method described here may be equally applicable to the sequencing of DNA fragments labeled at the 3′ end.

At present, we do not know the reactions which produce the initial labilization and removal of the various heterocyclic bases in hot aqueous piperidine. Scissions at A and G sites probably involve initial nucleophilic attack by piperidine on C-8 in the purine moiety, with opening of the imidazole ring. A process of this type has been documented by Albert & Brown (1954) for the reaction of 2-methylthiopurine with methylamine or dimethylamine at high temperatures. The reactions occurring at the pyrimidine bases may be related to the attack of hydroxide ion on deoxycytidine 5′-phosphate and deoxythymidine 5′-phosphate at high temperature, which was explored by Jones et al. (1966) and which has been developed into a variant (the A + C reaction) in the established chemical cleavage procedure (Maxam & Gilbert, 1980).

Nucleophilic attack on the bases as the first step to ultimate chain scission would explain our observation that with rising salt concentration the overall rate of breakdown of the DNA increases. Attack of uncharged piperidine on the uncharged bases A and C to give dipolar initial adducts should be facilitated by high ionic strength. This will also be the case with nucleophilic attack by piperidine at those guanine and thymine moieties which remain uncharged. The relative decrease in susceptibility of G sites to cleavage which is seen at high salt (2.5 M) probably reflects the greater extent of ionization of guanine moieties in the highly basic medium (pH $\simeq$12) as the salt concentration increases; few guanines will then remain uncharged and thus liable to nucleophilic attack. It is still unexplained, however, why increasing salt concentration is not accompanied by a similar relative reduction in cleavage at the thymine bases.

Except for the interpretation of the first several bands, the method described in this report approaches, but probably does not attain, the accuracy of the established four-lane methods, viz., the chemical cleavage method of Maxam & Gilbert

(1977) and the dideoxyribonucleotide terminator method of Sanger et al. (1977). For reliable sequencing, the one-lane method will be more dependent on confirmation by sequencing of the complementary strand. The method should be particularly useful in routine sequencing applications, as in identification of known sequences, confirmation of known sequences, or in comparative sequencing of mutant DNAs. Work to extend and refine this approach is proceeding in our laboratory.

ACKNOWLEDGMENTS

We are indebted to Lee F. Johnson, John N. Reeve, Elio F. Vanin, Sylvia Perryman, Cynthia Rossano, Rell W. Ambrose, and Paul Anziano for generous help and advice.

Registry No. Piperidine, 110-89-4.

REFERENCES

Albert, A., & Brown, D. J. (1954) *J. Chem. Soc.*, 2060-2071.

Brownlee, G. G., & Sanger, F. (1967) *J. Mol. Biol. 23*, 337-353.

Jones, A. S., Mian, A. M., & Walker, R. T. (1966) *J. Chem. Soc. C*, 1784-1786.

Maxam, A. L., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560-564.

Maxam, A. L., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499-560.

Pless, R. C., & Bessman, M. J. (1983) *Biochemistry 22*, 4905-4915.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.

Sealey, P. G., & Southern, E. M. (1982) in *Gel Electrophoresis of Nucleic Acids: A Practical Approach* (Rickwood, R., & Hames, B. D., Eds.) pp 52-53, IRL Press Limited, Oxford and Washington, DC.

Southern, E. M. (1979) *Anal. Biochem. 100*, 319-323.

Sutcliffe, J. G. (1979) *Cold Spring Harbor Symp. Quant. Biol. 43*, 77-90.

Internal Motions in B- and Z-Form Poly(dG-dC)·Poly(dG-dC): ^{1}H NMR Relaxation Studies†

Peter A. Mirau,‡ Ronald W. Behling, and David R. Kearns*

Department of Chemistry, University of California—San Diego, La Jolla, California 92093

Received December 31, 1984; Revised Manuscript Received May 31, 1985

ABSTRACT: Proton NMR relaxation measurements are used to compare the molecular dynamics of 60 base pair duplexes of B- and Z-form poly(dG-dC)·poly(dG-dC). The relaxation rates of the exchangeable guanine imino protons (G_{im}) in H_2O and in 90% D_2O show that below 20 °C spin–lattice relaxation is exclusively from proton–proton magnetic dipolar interactions while proton–nitrogen interactions contribute about 30% to the spin–spin relaxation. The observation that the spin–lattice relaxation is nonexponential and that the initial spin–lattice relaxation rate of the G_{im}, G-H8 and C-H6 protons depends on the selectivity of the exciting pulse shows that spin–diffusion dominates the spin–lattice relaxation. The relaxation rates of the G_{im}, C-H5, and C-H6 in B- and Z-form poly(dG-dC)·poly(dG-dC) cannot be explained by assuming the DNA behaves as a rigid rod. The data can be fit by assuming large-amplitude out of plane motions (±30–40°, τ = 1–100 ns) and fast, large-amplitude local torsional motions (±25–90°, τ = 0.1–1.5 ns) in addition to collective torsional motions. The results for the B and Z forms show that the rapid internal motions are similar and large in both conformations although backbone motions are slightly slower, or of lower amplitude, in Z DNA. At high temperatures (>60 °C), imino proton exchange with solvent dominates the spin–lattice relaxation of B-form poly(dG-dC)·poly(dG-dC), but in the Z form no exchange contribution (<2 s^{-1}) is observed at temperatures as high as 85 °C. Conformational fluctuations that expose the imino protons to the solvent are strikingly different in the B and Z forms. The results obtained here are compared with those previously reported for poly(dA-dT)·poly(dA-dT).

There is considerable interest in the sequence-dependent properties of DNA and the possible role that conformational heterogeneity plays in the biological functions of DNA and the sequence specific recognition by proteins. Recent X-ray diffraction studies of single crystals of DNA oligonucleotides have revealed significant sequence effects on DNA conformation (Izatt et al., 1971; Williams, 1972; Wang et al., 1979, 1981; Wing et al., 1980) and suggested possible roles in the interaction of DNA with proteins (Dickerson & Drew, 1981; Dickerson, 1983; Rich et al., 1984). Transient fluctuations in the local conformation of the DNA may also play a role in moderating DNA–protein interactions and, if large enough, could significantly diminish the importance of sequence effects on the time-averaged conformations of the DNA. In addition to possible biochemical importance, an evaluation of internal motions in DNA is crucial to current efforts to use 1D and 2D NMR relaxation techniques to determine the structures of DNAs in solution (Early et al., 1980b; Tjerneld, 1982; Feigon et al., 1983a,b; Haasnoot et al., 1983; Keepers & James, 1984).

NMR studies have already been used to characterize the internal motions in the backbone of DNA and RNA. Phos-

† This work was supported by the National Science Foundation (Grant PCM-8303374 to D.R.K.).

‡ Present address: AT&T Bell Laboratories, Murray Hill, NJ 07974.

0006-2960/85/0424-6200$01.50/0