Synthesis and Thermal Melting Behavior of Oligomer·Polymer Complexes Containing Defined Lengths of Mismatched dA·dG and dG·dG Nucleotides†

Jerry B. Dodgson‡ and Robert D. Wells*

ABSTRACT: Model DNA polymers containing heteroduplex regions of defined sequence and size were synthesized using polynucleotide phosphorylase and calf thymus terminal transferase. Heteroduplexes were of the form $(dG)_n$·$d(C_{12}A_mC_{\bar{x}})$, where m = 1-6, and $(dG)_n$·$d(C_{10}G_mC_{\bar{x}})$, where m = 1 and 3-5. Thermal melting studies of the model DNAs indicated that the heteroduplex regions did not disrupt the cooperative interaction between the flanking regions of dG·dC base pairs. Thus, it is possible that the heteroduplex nucleotides are accommodated in a stacked helical structure.

Previous studies from this laboratory (Chan and Wells, 1974; Chan et al., 1977) demonstrated that the *lac* repressor binding capacity of λp*lac* DNA was maximally reduced by as few as 2-5 nicks by single-strand specific nucleases, whereas approximately 300 nicks by any of three nonspecific nicking agents were necessary to give the same effect. The single-strand specific nucleases tested were those from *Aspergillus oryzae* (S_1 nuclease (Ando, 1966)) and from mung bean sprouts (Johnson and Laskowski, 1968). Recent mapping experiments demonstrated the presence of a specific single-strand nick near, but not at, the *lac* repressor and/or RNA polymerase binding sites (Chan et al., 1977). In an effort to better understand the DNA structural feature recognized by these enzymes, we have determined the nuclease susceptibility of several types of model DNAs containing defined regions of mismatched nucleotides (Dodgson and Wells, 1977).

A series of heteroduplex DNA molecules containing regions of dA·dG or dG·dG mismatches of defined length was prepared. These DNAs were enzymatically synthesized such that each preparation contained from one to six adjacent noncomplementary bases of known sequence. Each heteroduplex region was flanked by blocks of duplex dG·dC base pairs.

The general scheme of the synthesis of defined dA·dG and dG·dG heteroduplex regions is outlined in Figure 1. The protocol involved the use of the PNPase[1] of *E. coli* B and calf thymus terminal transferase. Reversed phase column chromatography on RPC-5 resin and the use of analytical 20% polyacrylamide gel electrophoresis facilitated the isolation and characterization of the individual deoxynucleotide oligomers. Thermal denaturation studies were performed on the heteroduplex DNAs in an effort to understand the extent of structural disruption caused by the mismatched nucleotides.

Materials and Methods

Nucleotides. Nonradioactive dADP, [^{3}H]dADP, [^{3}H]-dGDP, and [^{14}C]dCTP were purchased from Schwarz/Mann. dGDP was from Calbiochem. [α-^{32}P]dCTP and [^{3}H]dCTP were purchased from New England Nuclear. All nucleotides were greater than 80% pure, as assayed by paper chromatography, except dADP which was 70% pure (30% was dAMP).

Nucleic Acids. $(dG)_n$ and $(dC)_n$ were prepared and characterized as described (Wells et al., 1970). $(dC)_m$ oligomers were purified by RPC-5 column chromatography of a pancreatic DNase digest of $(dC)_n$ as described (Burd et al., 1975b). Oligomers less than 20 nucleotides in length were at least 90% pure. Those more than 20-nucleotides long were greater than 70% pure with the remaining 30% symmetrically distributed among the adjacent nucleotides ($m \pm 1$), except for $(dC)_{22}$ which was 70% $(dC)_{22}$ and 30% $(dC)_{21}$.

Enzymes. Polynucleotide phosphorylase from *E. coli* B was purified as described (Gillam and Smith, 1974), except that the final Sephadex G-100 gel filtration was omitted, and an additional purification step was performed on DEAE-Sephadex A25 as described (Kimhi and Littauer, 1968) followed by the repetition of the Sephadex G-200 gel filtration step (Gillam and Smith, 1974). One unit of PNPase catalyzed the polymerization of 1 nmol of dADP in 1 min when assayed as described. The specific activity of the enzyme was 60 units/mg of protein. Protein concentration was measured as described (Murphy and Kies, 1960). The enzyme was tested for DNase contamination by incubating [^{3}H]$(dC)_{10}$ and dADP with 4 units of PNPase in a 1-mL mock dADP addition reaction as described below. After 2 h, less than 5% of the radioactivity failed to bind to RPC-5 in 0.2 M KCl. Oligomers smaller than approximately $(dC)_7$ would not bind to such a column. Calf thymus terminal transferase was a generous gift from Dr. Robert Ratliff, Los Alamos Scientific Laboratory, Los Alamos, N.M. and has been described (Hayes et al., 1966). Micrococcal nuclease (*S. aureus*) was obtained from Mann Research Laboratories. Pancreatic deoxyribonuclease I and spleen phosphodiesterase were from Worthington Biochemical Corp.

† From the Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706. This work was supported by funds from the National Science Foundation (Grant BMS74-21420) and the National Cancer Institute (CM-12275).

‡ Present address: Department of Chemistry, California Institute of Technology, Pasadena, Calif. 91109. Supported by a predoctoral fellowship from the National Science Foundation, the Wharton Fellowship from the Department of Biochemistry, and a WARF Fellowship from the University of Wisconsin.

[1] Abbreviations and nomenclature: The subscript n designates a long-chain polymer equal to or greater than 500 nucleotides in length. m is used to designate a short oligomer or stretch of nucleotides of a single defined length. $\bar{x}$ is the average length of a stretch of nucleotides whose length is distributed from, for example, $\bar{x} - 7$ to $\bar{x} + 7$. PNPase, polynucleotide phosphorylase of *E. coli* B; DEAE, diethylaminoethyl; EDTA, (ethylenedinitrilo)tetraacetic acid; Tris, 2-amino-2-hydroxymethyl-1,3-propanediol.

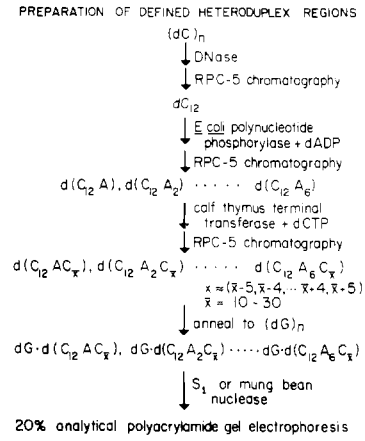

FIGURE 1: Scheme of synthesis of heteroduplex DNAs containing defined regions of dA·dG mismatched base oppositions. The procedures are described in detail in the text.

Other Materials. RPC-5 chromatographic resin was obtained from Miles Laboratories. Acrylamide and other materials were as described (Burd and Wells, 1974).

PNPase Reaction of $(dC)_{12}$ + dADP. 60 A_{270} units of $(dC)_{12}$ was incubated with 40 units of PNPase in a 10-mL reaction mixture containing 0.1 M Tris-HCl, pH 8.5, 0.01 M β-mercaptoethanol, 1.5 mM dADP, 12.5 mM $MnCl_2$, 0.2 M NaCl, and 15 μCi/mL of [^{3}H]dADP. After 3 h at 37 °C, the reaction was terminated by addition of 0.25 mmol of NaEDTA and heated to 95 °C for 3 min to inactivate the enzyme.

PNPase Reaction of $(dC)_{10}$ + dGDP. 63 A_{270} units of $(dC)_{10}$ was incubated with 80 units of PNPase in a 10-mL reaction mixture similar to the dADP reaction, except that 1.5 mM dGDP and 0.50 μCi/mL of [^{3}H]dGDP replaced the dADP, and the NaCl concentration was 0.05 M. After 4 h at 37 °C, the reaction was terminated with EDTA and the enzyme was heat inactivated as above.

High Salt PNPase Reaction of $(dC)_{10}$ + dGDP. In order to isolate $d(C_{10}G)$, a 9-mL reaction mixture containing 25 A_{270} units of $(dC)_{10}$ and 45 units of PNPase was incubated as above, except that the NaCl concentration was 0.45 M. The reaction was terminated after 105 min at 37 °C.

Digestion of Oligomers to 3′-Deoxyribonucleoside Monophosphates. Approximately 0.1–0.2 optical density unit of $d(C_{10}G_m)$ oligomer was incubated with 50 μg of micrococcal nuclease (*S. aureus*) in 0.1 mL of 0.01 M Tris-HCl, pH 8.5, 0.01 M $CaCl_2$ for about 18 h at 37 °C. The incubation mixtures contained 1 mM 5′-dGMP and 0.74 mM deoxyguanosine to facilitate recovery of labeled nucleotides and nucleoside. The reaction mixture was then adjusted to pH 7 with 5 μmol of potassium phosphate, pH 6.5, and 0.1 unit of spleen phosphodiesterase was added followed by further incubation at 37 °C for 2 h. The reaction was terminated with EDTA and heating as above. Approximately 50 μL of the reaction mixture was chromatographed on Whatman No. 1 paper in a descending isobutyric acid–ammonia–water system (57:4:39, respectively). In this system, 3′-dGMP migrated slightly ahead of the 5′-dGMP marker, but it was well separated from deoxyguanosine. Chromatograms were dried and cut into strips for liquid scintillation counting. Less than 0.5% of the ^{3}H radioactivity from the digestion of $d(C_{10}G_m)$ remained at the origin with at least 99% running as either 3′-dGMP or GdR. In a control digestion of [^{3}H]$(dG)_n$·$(dC)_n$ the amount of isotope in GdR was less than 3% of that in 3′-dGMP. Correction for this amount of phosphorolysis does not significantly alter the results.

Terminal Transferase Reactions. Conditions for the terminal transferase reactions were described (Burd and Wells, 1974). Primer concentrations varied from 60 to 600 μM (nucleotide) and dCTP varied from 70 to 700 μM. [^{3}H]-, [^{14}C]-, and [α-^{32}P]dCTP were used in different experiments generally at concentrations such that the counts per minute in dCTP were one to two times those in the primer. Incubations were for 2–4 h at 37 °C with 7–17 × 10^3 units of enzyme per mL, and generally greater than 80% of the dCTP was incorporated. The reactions were terminated with EDTA, heated to 95 °C for 3–5 min, and layered onto an RPC-5 column (1.6 × 25 cm). Unincorporated dCTP was washed off the column with 0.2 M KCl. (All RPC-5 buffers contain 0.01 M Tris-HCl, pH 7.9, 0.02% sodium azide.) Any primer that was not utilized was eluted with 0.45 M KCl (generally there was none detectable (<5%)). Product oligonucleotide was eluted from the column with 0.8 or 1.0 M KCl or alternatively with a 600-mL gradient from 0.45 M to 1.0 M KCl. Product oligonucleotides were pooled, concentrated, and dialyzed repeatedly, and their chain length distribution was determined by analytical polyacrylamide gel electrophoresis as described (Burd and Wells, 1974).

Annealing of Polymer·Oligomer Mixtures. Heteroduplex oligomers ($d(C_{12}A_mC_{\bar{x}})$ or $d(C_{10}G_mC_{\bar{x}})$) or homoduplex oligomers ($(dC)_m$) were combined with $(dG)_n$ in approximately equimolar ratios ($(dG)_n$ was kept in slight excess) in 0.1 M NaOH, and the mixture was dialyzed (without agitation) overnight into 0.01 M Tris-HCl, 0.1 mM NaEDTA adjusted to pH 10 with NaOH. The polymer·oligomer duplexes were then dialyzed into 0.01 M Tris-HCl, pH 7.9, 0.1 mM EDTA, or 0.01 M Na_2HPO_4, 0.1 mM NaEDTA, pH 8.8.

Thermal-Denaturation Studies. Helix–coil transitions were monitored with a Gilford 2000 spectrometer as described previously (Burd et al., 1975b). All samples were dialyzed extensively into 0.01 M Na_2HPO_4 and 0.1 mM NaEDTA and had a final pH of 8.8 ([Na^+] = 20 mM). All melting transitions were reversible and the t_M values are generally ±1 °C. Different preparations of the same oligomer·polymer complexes resulted in the same t_M values even when different batches of $(dG)_n$ were used. All thermal-denaturation curves were measured at λ_{max} which varied between 250 and 255 nm. In some cases, low levels of drift in absorbance were observed in the thermal-denaturation profile, presumably due to absorbance transitions of $(dG)_n$ self-structure. These were not subtracted out in the curves presented but were ignored when estimating the t_M. In any case, where such an estimate was not unambiguous, the difference in calculated t_M was not significant.

Other Techniques. Twenty percent polyacrylamide gel electrophoresis was described (Burd and Wells, 1974). Oligonucleotide purity was judged by electrophoresis of about 0.1 A_{270} unit of oligomer, and oligomer length was judged by coelectrophoresis with a pancreatic DNase digest of $(dC)_n$ or by comparison of R_f (relative to bromophenol blue or a dC oligomer) to a curve standardized with the $(dC)_n$ digest. All lengths were determined relative to dC oligomer standards described previously (Burd and Wells, 1974) (gift of J. F. Burd). Spectral analyses and other techniques were as described (Wells et al., 1970).

Results

The Synthesis of Defined Heteroduplex DNAs. The scheme of synthesis of dA·dG heteroduplex DNAs is shown in Figure 1. In these heteroduplex DNAs, one strand is comprised of poly(deoxyguanosine). The opposite strand is composed of

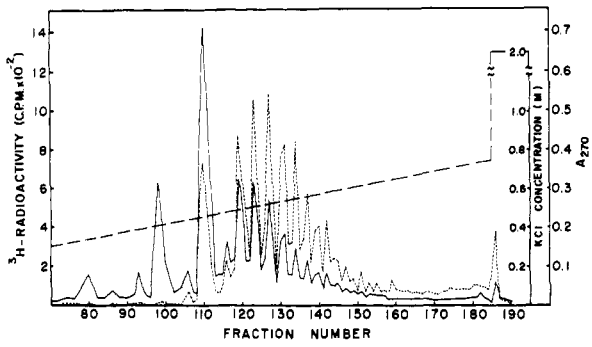

FIGURE 2: RPC-5 elution profile of the product of dADP addition onto $(dC)_{12}$. The PNPase reaction mixture (Methods) was diluted twofold with H_2O and applied to a 1.6 × 90 cm RPC-5 column packed in 0.2 M KCl. The column was washed with approximately 300 mL of 0.2 M KCl, and then a 1.5-L gradient from 0.2 to 0.75 M KCl was applied to the column, followed by a wash with 2.0 M KCl. Eleven-milliliter fractions were collected, their absorbance was read at 270 nm, and the radioactivity in a 20-μL sample of each fraction was measured. (—) Absorbance; (- - -) radioactivity; (– –) KCl concentration.

blocks of complementary dC residues flanking a region of noncomplementary dA residues from one to six nucleotides in length. Upon annealing the strands, dA·dG heteroduplex regions one to six base pairs long are obtained. Defined dG·dG heteroduplex regions one to five nucleotides long have been synthesized similarly.

The block of dC residues on the 5′ end of the heteroduplex region was supplied by a dC oligomer purified from a pancreatic DNase digest of poly(deoxycytidine), in this case $(dC)_{12}$. This oligomer was used as a primer for the polynucleotide phosphorylase of *Escherichia coli* B with dADP as the substrate nucleotide. In such reactions, the PNPase adds from one to a few nucleotides to the 3′ end of the primer (Gillam et al., 1974, 1975a; Gillam and Smith, 1974). This allowed the formation of $d(C_{12}A_m)$ oligomers in high yields where m = 1-6. The block of dC residues on the 3′ end of the heteroduplex oligomers was added with calf thymus terminal transferase. This enzyme uses dCTP as a substrate to extend the $d(C_{12}A_m)$ oligomers from the 3′ hydroxyl end to give a group of oligomers of the form $d(C_{12}A_mC_{\bar{x}})$ where x may range, for example, from $\bar{x} - 5$ to $\bar{x} + 5$. The product oligomer of the desired size distribution was then purified and annealed to $(dG)_n$ to give defined dA·dG heteroduplex regions which were m nucleotides in length. The identical scheme was used to synthesize defined dG·dG heteroduplex regions one and three to five nucleotides in length, except that $(dC)_{10}$ was used with dGDP in the PNPase reaction.

The repeated use of RPC-5 chromatography to purify oligo(deoxynucleotides) was crucial to these syntheses. This resin (Pearson et al., 1971; Burd et al., 1975a) allows individual oligomers to be purified in the high yields necessary to synthesize the heteroduplex DNAs of carefully defined composition in the heteroduplex region. Analytical 20% polyacrylamide gel electrophoresis (Burd and Wells, 1974) was also central to the characterization of the oligomers at each step.

Synthesis of $d(C_{12}A_m)$. The use of *E. coli* B polynucleotide phosphorylase for the controlled addition of single deoxynucleotide residues to oligo(deoxynucleotide) primers using deoxynucleoside diphosphates was examined extensively by Smith and co-workers (Gillam et al., 1974, 1975a; Gillam and Smith, 1974). They found that the major product of such reactions was the single nucleotide addition product with smaller or negligible yields of multiple addition products. The products of the PNPase reaction of [^{3}H]dADP with $(dC)_{12}$ primer as

TABLE I: Chain Length Determination of $d(C_{12}A_m)$ Pools from RPC-5 Chromatography.

Oligomer	mol of nucleotide per mol [^{3}H]dA	Theoretical mol of nucleotide/mol of dA
$d(C_{11}A)$	12.6	12.0
$d(C_{12}A)$	(13.0)[a]	13.0
$d(C_{12}A_2)$	7.3	7.0
$d(C_{12}A_3)$	5.2	5.0
$d(C_{12}A_4)$	4.2	4.0
$d(C_{12}A_5)$	3.6	3.4
$d(C_{12}A_6)$	3.1	3.0

[a] The parentheses indicate that the moles of nucleotide to moles of dA ratio was set to 13.0 in order to calculate the extinction coefficient (6.9×10^3) used to calculate the ratios for the other oligomers.

resolved by RPC-5 chromatography are shown in Figure 2. The first major peak to elute was shown to be unreacted $(dC)_{12}$ by gel electrophoresis and by the relative absence of radioactivity. This peak was followed by peaks characterized as $(dC_{11}A)$, $d(C_{12}A)$, $d(C_{12}A_2)$, $d(C_{12}A_3)$, etc. Note that the major product is $d(C_{12}A)$, in agreement with Gillam et al. (1974). We do, however, find considerably greater yields of multiple addition products. This is due partially to the lower NaCl concentration (0.2 M) used in our PNPase reaction, but analytical scale reactions performed in the presence of 0.3 M NaCl also show high levels of multiple addition. This may be related to the nature of the $(dC)_{12}$ primer. Note also the unusually low yield of $d(C_{12}A_2)$. This has been observed repeatedly. The presence of a minor $d(C_{11}A)$ peak was also observed. This results from the phosphorolytic action of PNPase on $(dC)_{12}$ and subsequent dA addition. Smaller amounts of $d(C_{11}A_m)$ products may be concealed under peaks corresponding to $d(C_{12}A_{m-1})$, but in no case would the contamination be expected to exceed 5%.

The assignments of product composition suggested by the affinity of the products for RPC-5 and the known behavior of PNPase were confirmed by analytical polyacrylamide gel electrophoresis and the specific radioactivity of the purified oligomers. All oligomers were greater than 90% pure, as judged by gel electrophoresis (data not shown). Using the specific activity of the [^{3}H]dADP, it was possible to calculate an extinction coefficient of 6.9×10^3 at the λ_{max} (267) for $d(C_{12}A)$. Using this extinction coefficient, the ratio of moles of nucleotide to moles of [^{3}H]dA residues in the other oligomers was calculated. The excellent agreement with the theoretically expected ratios (Table I) confirms the oligomer assignments.

Synthesis of $d(C_{10}G_m)$. The RPC-5 profile of the products of the dGDP addition reaction is shown in Figure 3. Following the unreacted $(dC)_{10}$ peak, there are three major product oligonucleotide peaks. In contrast with the dADP reaction profile, the radioactivity to optical density ratio decreases in the later-eluting peaks. Upon isolation of the individual peaks, the oligomers appeared to be more than 90% pure on 20% gels, except for the earliest-eluting peak which contained 10–15% of a more rapidly migrating component. However, the R_f's of the three main products were only slightly different, and they could not be completely resolved from each other on the gels.

In order to determine the composition of these three oligomers, they were digested to 3′-deoxynucleoside monophosphates and the 3′-hydroxyl-terminal deoxynucleoside

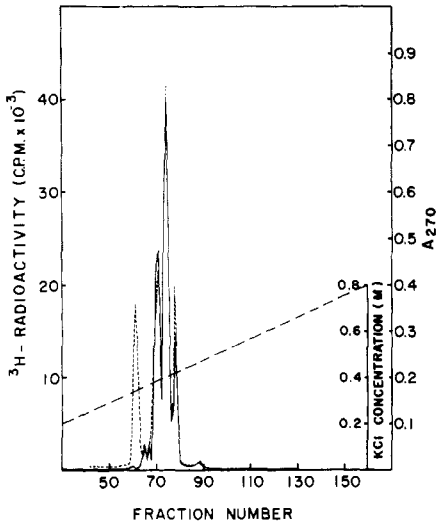

FIGURE 3: RPC-5 elution profile of the product of the low salt (0.05 M NaCl) dGDP addition to $(dC)_{10}$. The PNPase reaction mixture (Methods) was loaded onto a 1.6 × 90 cm RPC-5 column packed in 0.2 M KCl. The column was washed with 350 mL of 0.2 M KCl, followed by a 1.6-L gradient from 0.2 to 0.8 M KCl. Thirteen-milliliter fractions were collected and assayed for their absorbance at 270 nm and the ^{3}H radioactivity in a 50-μL sample. (- - -) Absorbance; (—) radioactivity; (– –) KCl concentration.

TABLE II: Chain Length Determination of the dG Region of $d(C_{10}G_m)$ Oligomers from Nucleotide to Nucleoside Ratios.

Oligomer	Molarity of KCl necessary for elution from RPC-5	3′-dGMP (%)[a]	GdR (%)[a]	Calcd ratio of 3′-dGMP to GdR	Theor ratio of 3′-dGMP to GdR
$d(C_{10}G_3)$	0.42	67	33	2.0	2.0
$d(C_{10}G_4)$	0.40	73	27	2.7	3.0
$d(C_{10}G_5)$	0.39	79	21	3.8	4.0

[a] One hundred percent corresponds to 18 000–40 000 cpm.

(Methods). By comparing the amounts of radioactivity in the 3′-dGMP peak relative to that in the GdR peak (only the deoxyguanosine residues are labeled), the length of the deoxyguanosine stretch can be determined. As can be seen from the results in Table II, the peaks, in order of their elution from RPC-5, are $d(C_{10}G_5)$, $d(C_{10}G_4)$, and $d(C_{10}G_3)$. While this was in general agreement with the radioactivity/OD ratios and the gel mobilities, it was unexpected, since single-stranded oligomers previously had eluted from RPC-5 in order of increasing size. This behavior probably is due to intramolecular hydrogen bonding between the dG stretch and the dC blocks in the individual oligomers, at least for those with more than 3 dG residues. In this way, the addition of one dG residue to one of the oligomers (i.e., going from $d(C_{10}G_3)$ to $d(C_{10}G_4)$) replaces one single-stranded nucleotide with two base-paired nucleotides. As previously shown (Hardies and Wells, 1976), double-stranded deoxynucleotides bind to RPC-5 much less tightly than single-stranded deoxynucleotides. This apparently explains the inversion of the expected elution pattern of the $d(C_{10}G_m)$ oligomers.

It is not known why the PNPase reaction in this case produces multiple rather than single nucleotide addition products, but it may also be related to the base-pairing properties of these oligonucleotides. We have tested a variety of salt conditions

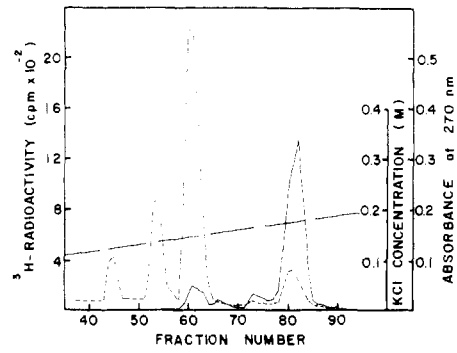

FIGURE 4: RPC-5 elution profile of the product of the high salt (0.45 M NaCl) dGDP addition to $(dC)_{10}$. The high salt PNPase reaction mixture (Methods) after a threefold dilution with H_2O was loaded onto a 1.5 × 23 cm RPC-5 column packed in 0.2 M KCl. The column was washed with 150-mL of 0.2 M KCl, followed by a 1-L gradient between 0.2 and 0.6 M KCl. Seven-milliliter fractions were collected and the absorbance at 270 nm and the ^{3}H radioactivity in 50 μL were measured. (- - -) Absorbance; (—) radioactivity; (– –) KCl concentration.

from 0 to 0.6 M NaCl, and in no case was a high percentage of $d(C_{10}G)$ produced. At higher salt conditions, the overall reaction is inhibited, and the multiple addition products are produced in low yields similar to that of $d(C_{10}G)$. Only about 10% of the $(dC)_{10}$ is converted to $d(C_{10}G)$ in any of these reactions. The RPC-5 elution profile of a reaction in 0.45 M NaCl is shown in Figure 4. On this short column, the $d(C_{10}G)$ was not fully resolved from $d(C_{10}G_{3-5})$. The earliest-eluting 70% of the major product peak was pooled and analyzed by gel electrophoresis and digestion to 3′-dNMPs. The latter procedure gave 32% 3′-dGMP and 68% GdR. This was consistent with a composition of about 50% $d(C_{10}G)$ and 50% $d(C_{10}G_2)$ or 75% $d(C_{10}G)$ and 25% $d(C_{10}G_{3-5})$. Gel electrophoresis favors the latter interpretation, since it indicates that the pool is composed of about 75% of a more rapidly migrating component and 25% of a component that comigrates with $d(C_{10}G_{3-5})$. The oligomer $d(C_{10}G_2)$ is apparently produced in only a very low yield in this system as was $d(C_{12}A_2)$ in the dADP addition system. These assignments of oligomer size are in agreement with the thermal stability of their complexes with poly(deoxyguanosine) after elongation with terminal transferase (see below).

Terminal Transferase Reactions. All of the oligonucleotides described above were efficient primers for the terminal transferase reaction which was used to add dC residues to the 3′-hydroxyl end of the oligomers. Generally, the dCTP was labeled with ^{14}C or α-^{32}P. (In one case where [^{14}C]dADP had been used in the PNPase reaction, [^{3}H]dCTP was used.) In all cases, about 80% of the dCTP was incorporated in 2 to 4 h at 37 °C (Methods). Therefore, by varying the dCTP concentration, the average size of the product oligomers can be varied. The distribution of product oligomers was slightly wider when $d(C_{10}G_m)$ primers were used than when $d(C_{12}A_m)$ primers were used. Also, by varying the salt (KCl) concentrations of the eluting buffer on the subsequent RPC-5 column (Methods), pools of defined oligomer size could be collected. Alternatively, a salt gradient could be used to elute the oligomers from RPC-5 and the fractions pooled to give the desired length distribution. A variety of such oligomer pools were synthesized and purified.

Table III lists some of the pools of the form $d(C_{12}A_mC_{\bar{x}})$, giving m, $\bar{x}$, x_{max}, x_{min}, and the average total oligomer length. Note that, although the radioactivity levels and absorbance of the oligomers can be used to measure $\bar{x}$, the more accurate

TABLE III: Description of Synthetic Oligomers of the Form $d(C_{12}A_mC_{\bar{x}})$.

m	$\bar{x}$[a]	x_{min}[b]	x_{max}[b]	Oligomer length $(12 + m + \bar{x})$
1	5	4	7	18
1	8	6	11	21
1	14	7	18	27
1	18	13	22	31
1	22	19	27	35
2	11	5	17	25
3	11	5	16	26
3	28	14	40	43
4	8	4	14	24
4	31	19	40	47
5	29	15	40	46
6	31	18	42	49

[a] Weight average length of the block of dC residues added to the 3'-hydroxyl of $d(C_{12}A_m)$, as judged by analytical gel electrophoresis (see text). [b] x_{min} and x_{max} correspond to the x values (length of the 3'-dC stretch) of the shortest and longest oligomers, respectively, in the distribution which gave a visible peak on analytical gels.

TABLE IV: Description of Synthetic Oligomers of the Form $d(C_{10}G_mC_{\bar{x}})$.

m	$\bar{x}$[a]	x_{min}[b]	x_{max}[b]	Oligomer length $(10 + m + \bar{x})$
1	25	12	40	36
3	17	10	30	30
4	9	7	11	23
4	20	11	36	34
5	23	11	37	38
5	42	35	53	57

[a] Weight average length of the block of dC residues added to the 3'-hydroxyl of $d(C_{10}G_m)$, as judged by analytical gel electrophoresis (see text). [b] Defined in Table III.

gel electrophoresis technique was used in all cases reported here. For these oligomers, this gives an error in $\bar{x}$ of about ±10%. In most cases, the approximate weight average oligomer length is within one nucleotide of the most prevalent oligomer.

Table IV gives similar information for oligomers of the form $d(C_{10}G_mC_{\bar{x}})$. In this case, the error in $\bar{x}$ may be slightly higher, since the $d(C_{10}G_m)$ series does not electrophorese with similar mobilities to the dC oligomers of equal length (see above). Therefore, the electrophoretic effect of the few dG nucleotides in the center of long dC nucleotide stretches is not certain, and this may influence the $\bar{x}$ measurements. The reported $\bar{x}$ measurements are probably 1–2 nucleotides low if they are in error at all.

Preparation of Polymer·Oligomer Duplexes. The $(dG)_n$·oligomer duplexes were prepared by slow dialysis of appropriate mixtures of nucleic acids from alkaline to neutral conditions (Wartell et al., 1974) (Methods). The alkaline conditions were apparently necessary to remove some inherent $(dG)_n$ self-structure that inhibited duplex formation. This procedure gave rise to stable duplexes, as judged by their hyperchromicity and resistance to single-strand specific nucleases (Dodgson and Wells, 1977). This does not seem to be the case for oligomers smaller than approximately $(dC)_{12}$ (our unpublished results). The polymer·oligomer complexes in this case are apparently less stable than the $(dG)_n$ self-structure. Oddly enough, although the thermal melting curves in low salt of these complexes are reversible, the complexes cannot be formed by heating and slow cooling in low salt. Apparently, $(dG)_n$ as purified has a very stable self-structure that must be eliminated with alkali treatment before dialysis into low salt conditions.

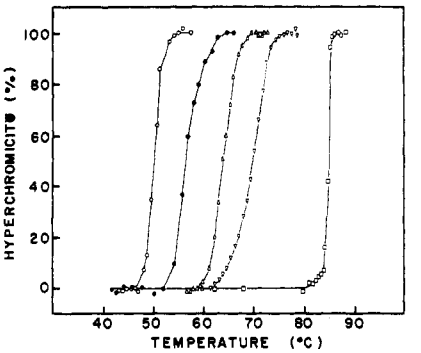

FIGURE 5: Melting transitions for homoduplexes of $(dG)_n$·$(dC)_m$. dC oligomers were annealed to $(dG)_n$ and absorbance–temperature profiles were determined as described (Methods). $(dG)_n$·$(dC)_{12}$ (○), $(dG)_n$·$(dC)_{15}$ (●), $(dG)_n$·$(dC)_{22}$ (△), $(dG)_n$·$(dC)_{30}$ (▽), and $(dG)_n$·$(dC)_n$ (□). The actual hyperchromicity increased from 8% for $(dG)_n$·$(dC)_{12}$ to 18% for $(dG)_n$·$(dC)_{30}$ and $(dG)_n$·$(dC)_n$.

In early experiments, the $(dG)_n$·oligomer complexes were purified by gel filtration on Sephadex G-75 as described previously (Burd and Wells, 1974) to remove any uncomplexed oligomer. Since less than 10% of the oligomer eluted after the void volume, this step was omitted in later preparations with no observable difference in the final products.

Thermal Melting of $(dG)_n$·$(dC)_m$ Homoduplex DNAs. As a control for the melting of defined heteroduplex DNA, the melting of analogous $(dG)_n$·$(dC)_m$ homoduplexes was studied. The preparation of these homoduplexes is described above and the thermal denaturation profiles for several complexes are shown in Figure 5. All profiles are similar except that the transitions of the oligomer·polymer complexes are somewhat broader than that of $(dG)_n$·$(dC)_n$.

As shown previously (Burd et al., 1975b), the t_M values of polymer·oligomer complexes are directly related to the inverse of the oligomer length. This is shown for the $(dG)_n$·$(dC)_m$ series in Figure 6. In this case, t_M oligomer = t_M polymer − 369 m^{-1}. This relationship is used for the calculation of the length of a homoduplex of equivalent thermal stability to a given heteroduplex.

Transitions of Oligomer Complexes Containing dA·dG Mismatches. The relative insensitivity of many of the above described heteroduplex samples to high levels of S_1 nuclease (Dodgson and Wells, 1977) suggested that the heteroduplex bases might be stacked in a double-helical form. Therefore, the thermal melting curves of such heteroduplex DNAs were examined in order to determine the effect of heteroduplex bases on the stability of the polymer·oligomer complexes.

The thermal denaturation profiles for a series of such heteroduplexes of the form $(dG)_n$·$d(C_{12}A_mC_{\bar{x}})$, where $m = 1$ and $\bar{x} = 5$, 8, 14, and 22, are shown in Figure 7A. Stability increases with increasing average oligomer length. It is notable that such complexes melt as a single cooperative unit, implying that the single heteroduplex base pair does not disrupt the cooperativity of the adjoining dC stretches. Note that transitions are not observed at the t_M of $(dG)_n$·$(dC)_{12}$ (51.5 °C), implying that the $(dC)_{12}$ nucleotide block is melting cooperatively with the rest of the oligomer. This is not due to the ex-

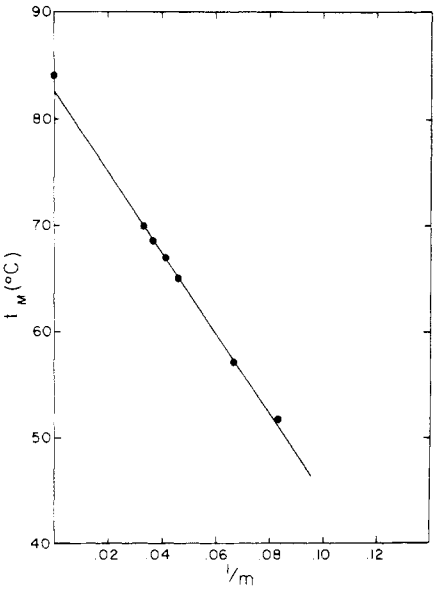

FIGURE 6: Thermal stability as a function of oligomer size for the duplexes of $(dG)_n \cdot (dC)_m$. t_M, determined as in Methods, was plotted vs. $1/m$ for $(dG)_n \cdot (dC)_m$ and for $(dG)_n \cdot (dC)_n$. The line drawn through the data points corresponds to the equation: t_M ($(dG)_n \cdot (dC)_m$) = 82.2 °C − (°C/m).

TABLE V: Thermal Melting of $(dG)_n \cdot d(C_{12}A_mC_{\bar{x}})$ Heteroduplexes.

m (nucleotides)	$\bar{x}$ (nucleotides)	t_M (°C)	Oligomer length of equally stable homoduplex (nucleotides)	Total heteroduplex length ($12 + m + \bar{x}$) (nucleotides)
1	5	56	14	18
1	8	60	17	21
1	14	66.5	24	27
1	18	68.5	27	31
1	22	70	30	35
2	11	63	19	25
3	11	58.5	16	26
3	28	70.5	32	43
4	8	55	14	24
4	31	71	33	47
5	29	69	28	46
6	31	69.5	29	49

clusion of the $(dC)_{12}$ block from the duplex by $(dG)_n$ self-structure, since each of the oligomers melts at a t_M much higher than that of either of the blocks of dG·dC base pairs alone.

The $(dG)_n \cdot d(C_{12}A_mC_{\bar{x}})$ heteroduplexes, where m is greater than one, also melt with uninterrupted cooperativity. Figure 7B shows the melting transitions for complexes of m = 1, 2, 3, and 4 with $\bar{x}$ = 14, 11, 11, and 8, respectively. Although the transitions are slightly broader for complexes with longer heteroduplex regions, these complexes also melt cooperatively throughout their length and not as two independent blocks of dC·dG base pairs separated by a heteroduplex block. These observations also apply to heteroduplexes 5 and 6 dA·dG base pairs in length (data not shown). The melting of $(dG)_n \cdot d(C_{12}A_mC_{\bar{x}})$ complexes with m = 3, 4, 5, and 6 and $\bar{x} \simeq 30$ gives transitions similar to those in Figure 7B, except that the t_M values are higher due to the greater $\bar{x}$ values (Table V). In all cases studied, the presence of a block of dA·dG heteroduplex

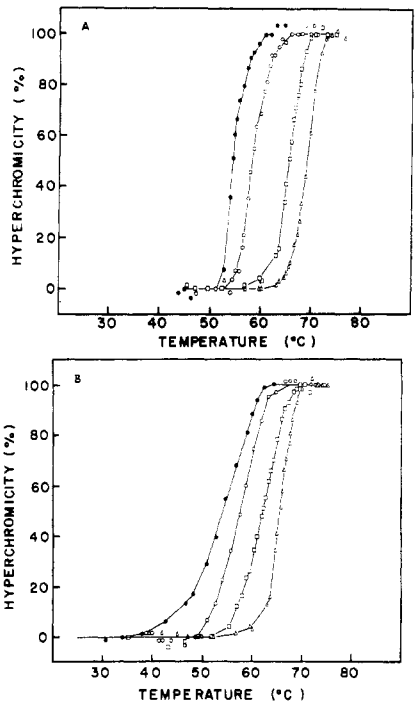

FIGURE 7: Melting transitions for heteroduplexes of the form $(dG)_n \cdot d(C_{12}A_mC_{\bar{x}})$. Heteroduplex DNAs, synthesized as described, were melted under standard conditions (Dodgson and Wells, 1976). One hundred percent hyperchromicity corresponds to an actual hyperchromic shift of from 10 to 18%. (A) Hyperchromicity is plotted vs. temperature for heteroduplexes of the form $(dG)_n \cdot d(C_{12}AC_{\bar{x}})$, where $\bar{x}$ = 5 (●), 8 (○), 14 (□), and 22 (△). (B) Hyperchromicity is plotted vs. temperature for melting transitions of DNAs of the form $(dG)_n \cdot d(C_{12}A_mC_{\bar{x}})$, where m = 1 (△) ($\bar{x}$ = 14); 2 (□) ($\bar{x}$ = 11); 3 (○) ($\bar{x}$ = 11); and 4 (●) ($\bar{x}$ = 8).

base pairs does not disrupt the cooperative interaction between the flanking blocks of dG·dC base pairs.

A summary of the thermal stability of a variety of DNA preparations containing dA·dG mismatched base pairs is given in Table V. For purposes of comparison, the length of a heteroduplex (i.e., value of m for $(dG)_n \cdot (dC)_m$) of equal stability (i.e., same t_M) as each of the heteroduplex DNAs is given. Recall that the t_M of these molecules varies linearly with m^{-1}. The presence of heteroduplex regions in these DNAs destabilized the complexes (i.e., decreased the t_M). For a given number of heteroduplex base oppositions, the difference in stability from the homoduplex situation may be expressed as the difference in the heteroduplex oligomer length from the oligomer length of the homoduplex of the same t_M. For example, the t_M of $(dG)_n \cdot d(C_{12}AC_{\bar{5}})$ is 56 °C which is the same as the calculated t_M of $(dG)_n \cdot (dC)_{14}$. The difference between the actual average oligomer length of 18 ($12 + m + \bar{x}$) and the length of the equally stable homoduplex is therefore 4 nucleotides. This is a convenient (although arbitrary) method for expressing the effect of heteroduplex regions on the stability of the polymer·oligomer complexes, since this difference is relatively constant for heteroduplexes containing the same number of heteroduplex dA·dG oppositions (i.e., the same m) regardless of total length (i.e., $12 + m + \bar{x}$). As can be seen from comparison of the two right-hand columns of Table V for m = 1, $(dG)_n \cdot d(C_{12}A_mC_{\bar{x}})$ is about 4 base pairs less stable than $(dG)_n \cdot (dC)_{(12+m+\bar{x})}$. For m = 2, 3, 4, 5, and 6, this difference in stability is about 6, 10, 12, 18, and 20 base pairs, respectively. Therefore, it appears that for each heteroduplex base opposition the overall stability of the resultant complex is reduced by an equivalent of about 3 to 4 base pairs.

TABLE VI: Thermal Melting of $(dG)_n \cdot d(C_{10}G_mC_{\bar{x}})$ Heteroduplexes.

m (nucleotides)	$\bar{x}$ (nucleotides)	t_M (°C)	Oligomer length of of equally stable homoduplex (nucleotides)	Total heteroduplex length $(10 + m + x)$ (nucleotides)
1	25	70	30	36
3	17	64	20	30
4	9	53	13	23
4	20	65	21	34
5	23	66	23	38
5	42	73	41	57

Transitions of Oligomer Complexes Containing dG·dG Mismatches. The melting profiles of the $(dG)_n \cdot d(C_{10}G_mC_{\bar{x}})$ DNAs are very similar to the $(dG)_n \cdot d(C_{12}A_mC_{\bar{x}})$ profiles shown above. That is, heteroduplex regions up to five dG·dG base oppositions in length do not disrupt the cooperative interaction between the two flanking blocks of dG·dC base pairs. The heteroduplex regions do destabilize the polymer·oligomer interactions, however, as seen in Table VI. The amount of destabilization again can be expressed as the difference between the total length of the heteroduplex and the length of an equally stable $(dG)_n \cdot (dC)_m$ homoduplex. The heteroduplex $(dG)_n \cdot d(C_{10}GC_{\overline{25}})$ is about 6 base pairs less stable than if it were a homoduplex of equal length ($(dG)_n \cdot (dC)_{36}$). This number is probably an overestimate, since this particular preparation is contaminated with about 25% of duplexes containing about three dG·dG mispairs. The $(dG)_n \cdot d(C_{10}G_mC_{\bar{x}})$ heteroduplexes in which m = 3, 4, and 5 are destabilized by about 10, 12, and 16 base pairs, respectively. Thus, it appears that for each dG·dG mismatched base opposition in the duplex molecules the resultant polymer·oligomer complexes are destabilized by an equivalent of about 3 dG·dC base pairs.

Discussion

The combined use of PNPase, terminal transferase, preparative RPC-5 column chromatography, and analytical 20% polyacrylamide gel electrophoresis has allowed the preparation and characterization of DNA duplexes containing dA·dG and dG·dG heteroduplex regions of defined length.

Absorbance–temperature transition studies were performed on the heteroduplex DNAs to obtain information on their structures. The thermal stability of model RNA polymers containing noncomplementary base oppositions was reviewed previously (Lomant and Fresco, 1975). Our results concerning the levels of destabilization by noncomplementary bases in model polymer DNA duplexes confirm and extend those found by previous workers (Gillam et al., 1975b). As noted above, the destabilization of a polymer·oligomer complex by a region of noncomplementary bases can be expressed by the difference between the oligomer length (in nucleotides) and the length of an equally stable homologous homoduplex. When expressed in this way, the observed destabilization was directly proportional to the length of the noncomplementary region; for each additional heteroduplex nucleotide, the complex stability was reduced equivalent to a length reduction of 3–4 base pairs. This does not indicate the disruption of three to four base pairs; it is merely a convenient method of quantitating the extent of destabilization, probably because it measures destabilization in units proportional to $(t_{M,\,polymer} - t_{M,\,oligomer})^{-1}$.

In fact, the actual t_M values of heteroduplex DNAs are only 0.5 to 4 °C less than the t_M values of analogous hypothetical duplexes containing dA·dT base pairs in place of every heteroduplex base opposition. For example, the t_M of $(dG)_n \cdot d(C_{12}AC_{\overline{22}})$ is 70 °C, whereas the t_M of $d(G_{\overline{22}}TG_{12}) \cdot d(C_{12}AC_{\overline{22}})$ would be 71 °C when calculated as described previously (Burd et al., 1975b). Therefore, the possibility of weak dA·dG and dG·dG base interactions cannot be eliminated. Adenine is known to interact with hypoxanthine bases in an $(A)_n \cdot 2(I)_n$ structure (Arnott and Bond, 1973), suggesting that a similar dA·dG interaction could occur in these heteroduplexes. Furthermore, the tendency of (dG) to form self-structure is well known (Wells and Wartell, 1974), suggesting the feasibility of dG·dG interactions.

Indeed, the rather weak susceptibility to single-strand specific nucleases of the model heteroduplex DNAs suggests that the heteroduplex bases may at least be stacked in some ordered structure, even if there is no specific interaction with the bases of the $(dG)_n$ strand. In general agreement with this interpretation is the observation that the heteroduplex region did not disrupt the cooperative interaction between the flanking blocks of dG·dC base pairs in any case. This implies that melting did not initiate at the heteroduplex region prior to the initiation of melting at the 5′ and 3′ ends of the oligomer as usual. This suggests, but does not prove, that the heteroduplex regions are not in a looped-out or bubble-type structure. Further physical studies of heteroduplex nucleotide interactions will be necessary to clarify this point.

Acknowledgments

We thank Mrs. J. E. Larson for preparation and characterization of DNA polymers, Drs. H. W. Chan and R. L. Ratliff for gifts of enzymes, and Dr. J. F. Burd for helpful discussions.

References

Ando, T. (1966), *Biochim. Biophys. Acta 114,* 158.
Arnott, S., and Bond, P. J. (1973), *Science 181,* 68.
Burd, J. F., Larson, J. E., and Wells, R. D. (1975a), *J. Biol. Chem. 250,* 6002.
Burd, J. F., Wartell, R. M., Dodgson, J. B., and Wells, R. D. (1975b), *J. Biol. Chem. 250,* 5109.
Burd, J. F., and Wells, R. D. (1974), *J. Biol. Chem. 249,* 7094.
Chan, H. W., Dodgson, J. B., and Wells, R. D. (1977), *Biochemistry 16* (preceding paper in this issue).
Chan, H. W., and Wells, R. D. (1974), *Nature (London) 252,* 205.
Dodgson, J. B., and Wells, R. D. (1977), *Biochemistry 16* (following paper in this issue).
Gillam, S., and Smith, M. (1974), *Nucleic Acids Res. 1,* 1631.
Gillam, S., Waterman, K., Doel, M., and Smith, M. (1974), *Nucleic Acids Res. 1,* 1649.
Gillam, S., Waterman, K., and Smith, M. (1975a), *Nucleic Acids Res. 2,* 613.
Gillam, S., Waterman, K., and Smith, M. (1975b), *Nucleic Acids Res. 2,* 625.
Hardies, S. C., and Wells, R. D. (1976), *Proc. Natl. Acad. Sci. U.S.A. 73,* 3117.
Hayes, F. N., Mitchell, V. E., Ratliff, R. L., Schwartz, A. W., and Williams, D. L. (1966), *Biochemistry 5,* 3625.
Johnson, P. H., and Laskowski, M., Sr. (1968), *J. Biol. Chem. 243,* 3421.

Kimhi, Y., and Littauer, U. Z. (1968), *J. Biol. Chem. 243,* 231.
Lomant, A. J., and Fresco, J. R. (1975), *Prog. Nucleic Acid Res. Mol. Biol. 15,* 185.
Murphy, J. B., and Kies, M. W. (1960), *Biochim. Biophys. Acta 45,* 382.
Pearson, R. L., Weiss, J. F., and Kelmers, A. D. (1971), *Biochim. Biophys. Acta 228,* 770.
Wartell, R. M., Larson, J. E., and Wells, R. D. (1974), *J. Biol. Chem. 249,* 6719.
Wells, R. D., Larson, J. E., Grant, R. C., Shortle, B. E., and Cantor, C. R. (1970), *J. Mol. Biol. 54,* 465.
Wells, R. D., and Wartell, R. M. (1974), *MTP Int. Rev. Sci.: Biochem., Ser. One, 1973-6,* 41.

Action of Single-Strand Specific Nucleases on Model DNA Heteroduplexes of Defined Size and Sequence†

Jerry B. Dodgson‡ and Robert D. Wells*

ABSTRACT: The sensitivity of the model DNAs containing dA·dG and dG·dG heteroduplex regions of defined length to S_1 and mung bean single-strand specific nucleases was tested by polyacrylamide gel electrophoretic analysis of the distribution of product oligonucleotides. Single-base mismatch heteroduplexes were extremely resistant to these nucleases, although low levels of cleavage at the heteroduplex nucleotide were observed at high nuclease concentrations. The nuclease sensitivity of dA·dG heteroduplex regions increased gradually as the length of the heteroduplex region increased from one to six nucleotides. The sensitivity of dG·dG heteroduplexes three to five nucleotides long was considerably greater than that of the single dG·dG mismatch.

The reason for our interest in determining the susceptibility of regions of nonpaired nucleotides to single-strand specific nucleases was presented in the preceding paper of this issue (Dodgson and Wells, 1977). Also, the preparation, characterization of, and t_M studies on defined heteroduplex DNAs were reported.

In other studies, single-strand specific nucleases have been used to probe for other types of irregular structural sites in double-stranded DNA. First, these nucleases specifically cleave some superhelical DNAs, perhaps due to transient or permanent unwinding of the helix in order to decrease the superhelical density (Beard et al., 1973; Wiegand et al., 1975; Germond et al., 1974; Kato et al., 1973; Mitra et al., 1976; Wang, 1974). These nuclease-sensitive sites may either be in specific regions of the superhelical DNA (Beard et al., 1973; Germond et al., 1974) or may occur at random throughout the molecule (Wiegand et al., 1975; Mitra et al., 1976; Bartok and Denhardt, 1976), depending on the type of DNA studied. Second, these nucleases specifically cleave heteroduplex DNAs containing insertion loops at the mutant site (Wiegand et al., 1975; Shenk et al., 1975); it has been proposed that the nucleases can, in certain cases, cleave heteroduplex molecules at sites of single-base change mutations (Shenk et al., 1975; Legerski et al., 1976). Third, S_1 nuclease cuts nicked DNA in the strand across from the break (Beard et al., 1973; Wiegand et al., 1975; Germond et al., 1974; Shenk et al., 1975; Shishido and Ando, 1975). It also weakly cleaves nucleotides (or small oligomers) from the ends of DNA duplexes (Shenk et al., 1975), an action termed "nibbling". Fourth, mung bean nuclease preferentially cleaves double-stranded DNA at regions containing high percentages of dA·dT base pairs (Chan et al., 1977; Johnson and Laskowski, 1970). While the single-strand specificity of these nucleases has suggested that sensitive sites in duplex DNA molecules must exist at least transiently as single-strand loops or bubbles, the actual site of recognition on duplex DNA is unclear, since many of the sensitive sites are of unknown structure.

On the basis of the S_1 and mung bean nuclease susceptibility studies on biosynthetic model DNAs with heteroduplex regions of defined length, we conclude that single-base change heteroduplexes can be cleaved only with extremely low efficiency, and that nuclease sensitivity increases substantially as heteroduplex size increases above three nucleotides in length.

Materials and Methods

Enzymes. S_1 nuclease was purified as described (Vogt, 1973) with the omission of the sulfo-Sephadex chromatography. Units are as described by Vogt, although in some cases the units were measured by the hydrolysis of radioactive λ DNA in the presence of calf thymus DNA carrier at 37 °C and converted to Vogt units by a previously determined correction factor. The specific activity of the S_1 nuclease was 120 000 units/mg of protein. Mung bean nuclease was purified as described (Chan, 1976) and was the generous gift of H. W. Chan.

Other Materials. Nucleotides, nucleic acids, and other materials were described previously (Dodgson and Wells, 1977).

Synthesis of Defined Heteroduplex and Corresponding Homoduplex DNAs. All materials and methods pertaining

† From the Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706. *Received December 16, 1976.* This work was supported by funds from the National Science Foundation (Grant BMS74-21420) and the National Cancer Institute (CM-12275).

‡ Present address: Department of Chemistry, California Institute of Technology, Pasadena, Calif. 91109. Supported by a predoctoral fellowship from the National Science Foundation, the Wharton Fellowship from the Department of Biochemistry, and a WARF Fellowship from the University of Wisconsin.