Articles

Branch Migration Mediated DNA Labeling and Cloning†

Robin S. Quartin, Magdalena Plewinska, and James G. Wetmur*
Department of Microbiology, Mount Sinai School of Medicine, New York, New York 10029
Received April 24, 1989; Revised Manuscript Received June 21, 1989

ABSTRACT: The sequence-dependent attachment (capture) of an oligodeoxynucleotide duplex containing a single-stranded tail can be mediated by branch migration into the end of a DNA molecule. Substitution of bromodeoxycytidine (BrdC) for deoxycytidine (dC) increased DNA–DNA hybrid stability. BrdC-containing oligodeoxynucleotides displaced dC-containing strands from duplexes with blunt ends or 3′-overhangs. In the later case the rate of displacement was of the same order of magnitude as DNA reassociation. A BrdC-containing displacer oligodeoxynucleotide was used for transient sequence-specific invasion at a particular *Pst*I site. The product was captured by use of T4 DNA ligase and a linker oligodeoxynucleotide. The capture rate was more than 300 times the rate observed for an unrelated *Pst*I site. This high degree of specificity required BrdC substitution. In addition, deliberate incorporation of an incorrect nucleotide into a displacer strand demonstrated that branch migration was terminated at a mismatch. A branched, BrdC-containing ligated product of a capture reaction was cloned and sequenced. The specific capture reaction may be used to label a particular DNA fragment prior to electrophoresis, to mark the specific fragment for affinity chromatography, or to facilitate cloning by introducing a new overhanging sequence compatible with a restriction endonuclease site in a cloning vector.

The cloning or recloning of a particular DNA fragment from a eukaryotic organism is normally carried out by the two-step process of library construction and library screening (Maniatis et al., 1982). Recloning experiments would be greatly facilitated if a library could be constructed that was initially highly enriched for the particular fragment. The determination of the size of a particular DNA restriction fragment in a mixture of fragments is normally carried out by the two-step process of gel electrophoresis and Southern blot (Southern, 1975) hybridization. A simpler procedure would be to label the particular fragment prior to electrophoresis. In this study, we present an experimental and theoretical analysis of single-stranded DNA branch migration (Lee et al., 1970) and demonstrate fragment-specific DNA labeling and cloning.

Experiments carried out more than 20 years ago showed that the substitution of bromine at position C5 of pyrimidines leads to increased duplex stability (Michelson et al., 1967). Radding et al. (1962) showed that dG·BrdC was a more thermally stable base pair than dG·dC. In another study, poly(dI)·poly(BrdC) had a melting temperature 26 °C higher than poly(dI)·poly(dC) (Inman & Baldwin, 1964), and it was further shown that poly(BrdC) displaced poly(dC) from a poly(dI)·poly(dC) duplex to form a new duplex with poly(dI) (Inman, 1964).

There are numerous examples of homologous strand displacement. The replication of the supercoiled, circular genome of mammalian mitochondria is initiated unidirectionally with the formation of a displacement loop (D-loop) (Robberson et al., 1972). D-Loop formation has also been demonstrated in vitro (Holloman et al., 1975). The uptake of a large fragment of homologous single-stranded DNA by a supercoiled plasmid is driven entropically by the energy stored in the negative supercoils (Beattie et al., 1977) and is analogous to a step in generalized recombination with branch migration (Radding et al., 1977). Release of the superhelical free energy by nicking the supercoil resulted in rapid loss of the single-stranded DNA due to a directed branch migration process driven by the difference in entropy between the constrained single strand of the D-loop and the released, unconstrained single-stranded DNA. A similar reaction has been demonstrated where a short strand was displaced from a linear duplex, presumably through a D-loop intermediate (Green & Tibbetts, 1981). Recently, RecA protein, with ATP as an energy source, was employed to reverse DNA branch migration reactions, permitting the formation of D-loops in linear molecules (Rigas et al., 1986).

R-Loop formation, which is favored in 70% formamide, results when an RNA strand hybridizes to one strand of a DNA duplex and displaces the other (Thomas et al., 1976). The reverse reaction is favored in aqueous solution. The displacement of the RNA in an RNA–DNA hybrid by single-stranded DNA is the basis of a homogeneous nucleic acid hybridization assay, which detects ribonuclease digestion products by chemiluminescence (Vary, 1987).

The branch migration procedure we have developed uses "displacer" bromodeoxycytidine-containing oligodeoxynucleotides and DNA ligase. This method leads to uptake of a displacer oligodeoxynucleotide and covalent attachment of a partially complementary "linker" oligodeoxynucleotide at the end of a particular DNA fragment containing a sequence complementary to the displacer oligodeoxynucleotide. This specific attachment may be used (A) to label a particular fragment for detection without blotting and subsequent hybridization, (B) to mark a particular fragment for affinity chromatography, or (C) to facilitate cloning by introducing a new 5′- or 3′-overhang compatible with a restriction endonuclease site in a cloning vector.

Materials and Methods

Oligodeoxynucleotides, Plasmids, and Enzymes. All oligodeoxynucleotides (Table I) were synthesized on an Applied Biosystems Model 380B DNA synthesizer using standard

† This research was supported in part by grants from the National Institutes of Health and Enzo Biochem, Inc.

0006-2960/89/0428-8676$01.50/0

phosphoramidite chemistry. Bromodeoxycytidine incorporation was accomplished using a 5-bromodeoxycytidine phosphoramidite monomer (ABN-Fisher). Purification steps were limited to hydrolysis of base-protecting groups and cleavage from the support with NH_4OH, evaporation, resuspension, and ethanol precipitation. All oligodeoxynucleotides were 5′-labeled with ^{32}P by using T4 polynucleotide kinase, subjected to polyacrylamide gel electrophoresis on 20% acrylamide–8 M urea gels, and visualized by autoradiography. A single oligodeoxynucleotide species of the correct size was routinely detected. Plasmid pALA-D is a pUC9 expression vector containing the cDNA sequence of human δ-aminolevulinate dehydratase (ALA-D), a heme biosynthetic enzyme (Wetmur et al., 1986). pALA-D and pUC19 were each propagated in *Escherichia coli* strain HB101 and purified by standard methods (Maniatis et al., 1982). Restriction endonucleases, T4 polynucleotide kinase, and T4 DNA ligase were obtained from and used as recommended by New England Biolabs.

Melting Temperature Analyses. Melting temperatures for oligodeoxynucleotides were determined as previously described (Quartin & Wetmur, 1989). Briefly, complementary oligodeoxynucleotides were mixed 1:1 in 6× SSC, pH 4.0 or 7.0, or in 1 M NaCl, 0.05 M borate, and 0.2 mM EDTA, pH 10. Solutions were heated at 0.3 °C/min in a 1-cm quartz cuvette in a Beckman Model 25 spectrophotometer equipped with water-jacketed cell holder. The temperature was monitored by using a thermistor attached to the cell holder. The concentration-dependent melting temperatures ($T_m = t_m + 273.16$) were calculated according to the method of Marky and Breslauer (1987). T_m values reported are reliable to ±1 °C.

Displacement Assay. A gel migration assay was used to monitor the displacement of 5′-^{32}P-labeled dC-containing oligodeoxynucleotides from unlabeled complementary oligodeoxynucleotides by their BrdC-containing analogues. Labeled dC-containing oligodeoxynucleotides were annealed to unlabeled dC-containing complementary strands at room temperature in 1 M NaCl at concentrations of 1 and 3 μg/mL, respectively, and then brought to 4 °C. BrdC-containing analogues of the labeled strands (at concentrations ranging from 3 to 400 μg/mL) were incubated with dC-containing duplexes as a function of time at various temperatures. Aliquots taken at each time point were stored at −20 °C prior to electrophoresis. Samples were loaded onto a 20% acrylamide gel in 2.5% Ficoll 400, and electrophoresis was performed at 4 °C at 400 V and 5–15 mA in 89 mM Tris-HCl, 89 mM borate, and 1 mM EDTA, pH 8 (TBE). The gels were dried, and the results were obtained by autoradiography.

Attachment (Capture) and Cloning Procedures. A 14-mer linker oligodeoxynucleotide was ^{32}P-labeled by using T4 polynucleotide kinase. The reaction with labeled ATP was followed by reaction with excess unlabeled ATP to assure that all linker strands contained a 5′-phosphate. The linker oligodeoxynucleotides were purified over spun columns (Maniatis et al., 1982) of Sephadex G-50.

(1) Capture Ligation. pALA-D was digested with *Pst*I and *Rsa*I (Figure 1A). Digested plasmid (20 μg/mL) was mixed with a 34-mer displacer oligodeoxynucleotide (6 μg/mL) and 5′-^{32}P-labeled 14-mer linker oligodeoxynucleotide (2 μg/mL) in a pH-reduced T4 DNA ligase buffer (50 mM Tris-HCl, pH 7.0, 1 mM ATP, 10 mM $MgCl_2$, 20 mM DTT, 50 μg/mL bovine serum albumin), incubated at 55 °C for 10 min, cooled to room temperature, and incubated at 16 °C for from 1 min–20 h in the presence of from 5 to 500 units/mL of T4 DNA ligase. All units of ligase are New England Biolabs units. Ligation products were examined on ethidium bromide containing, 1% agarose gels by UV illumination and by autoradiography.

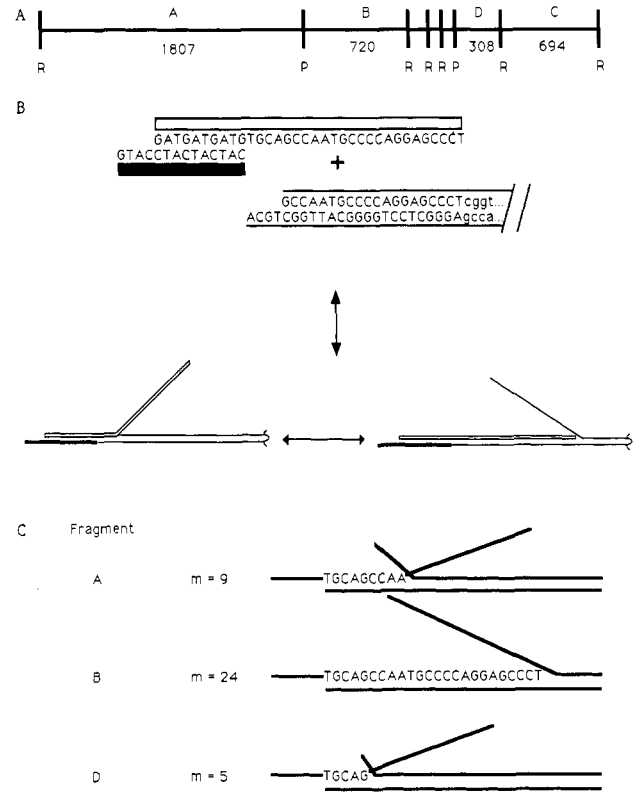

FIGURE 1: (A) Restriction map of pALA-D. R = *Rsa*I; P = *Pst*I. Fragments A–D are labeled above the line, with the nucleotide lengths indicated beneath. There is a single *Sma*I site in fragment D. (B) Branch migration of displacer (open rectangle), bound to linker (filled rectangle), into a recipient duplex with a four-base 3′-overhand (*Pst*I end of fragment B). Shown below is the conversion between the displacer–linker duplex bound to the 3′-overhang only (left) and following complete branch migration (right). (C) Maximum displacement with specific pALA-D fragments. m = the maximum number of base pairs that can be formed between the displacer and the complementary recipient strand.

(2) Ligation for Cloning. The capture ligation reaction was carried out with pALA-D digested with *Pst*I and *Sma*I (Figure 1A). The cleavage by *Sma*I in segment D prevented the pUC9 sequence in pALA-D from acting as a cloning vector. pUC19 was digested with *Pst*I and *Sph*I, both of which cut in the polylinker, added to 10 μL of the capture ligation mixture to a final concentration of 10 μg/mL, and incubated with 20000 units/mL T4 DNA ligase at 16 °C for 20 h. *E. coli* strain DH5α was used for the cloning experiments. Transformation (rubidium chloride) and identification of insert-containing β-galactosidase-negative clones were carried out according to previously described techniques (Maniatis et al., 1982). Double-stranded DNA sequencing was carried out with Sequenase (U.S. Biochemicals, Inc.), according to the protocol of the manufacturer, using both miniprep and large-scale alkaline lysis/CsCl-banded plasmid DNA.

Results

Displacement of dC-Containing Homologues by BrdC-Containing Oligodeoxynucleotides. The oligodeoxynucleotides used for melting temperature (t_m) analyses are listed in the first two sections of Table I. Table IIa presents the t_m values for dC-containing oligodeoxynucleotides and for their BrdC-containing analogues. The substitution of BrdC for dC leads to an increase in melting temperature, $\Delta T_m = T_m' - T_m$, where T_m and T_m' are the melting temperatures for dC- and BrdC-

Table I: Oligodeoxynucleotides[a]

	T_m Analysis
14-dC-S	ATG CAG CTA AGT CA
14-dC-A	TGA CTT AGC TGC AT
14-BrdC-S	ATG C̲AG C̲TA AGT C̲A
14-BrdC-A	TGA C̲TT AGC TGC̲ AT
	T_m and Displacement at Blunt Ends
12-dC-S	CAT GCA GCC CCA
12-dC-A	TGG GGC TGC ATG
12-BrdC-S	C̲AT GC̲A GC̲C C̲CA
	Displacement at Dangling Ends
16-dC-A	TGG GGC TGC ATG GCG T
16-BrdC-S	AC̲G C̲CA TGC̲ AGC̲ C̲CC A
	Attachment (Capture)
14-dC-L	CAT CAT CAT CCA TG
34-dC-D	GAT GAT GAT GTG CAG CCA ATG CCC CAG GAG CCC T
34-BrdC-D	GAT GAT GAT GTG C̲AG C̲CA ATG C̲CC C̲AG GAG C̲CC T
34-BrdC-D-E(10)	GAT GAT GAT GTG C̲AG C̲CA AAG C̲CC C̲AG GAG C̲CC T
34-BrdC-D-E(24)	GAT GAT GAT GTG C̲AG C̲CA ATG C̲CC C̲AG GAG C̲CC A

[a] Nomenclature: (oligodeoxynucleotide length)–(synthesis with dC or BrdC)–(S = sense; A = antisense; L = linker; D = displacer)–[E = error (mismatch) at stated position]. Abbreviation: C̲ = 5′-bromodeoxycytidine.

Table II: Bromodeoxycytidine Thermodynamics and Displacement Reactions

(a) Thermodynamics

oligodeoxynucleotides	t_m (°C) in 1 M Na^+, C = 6 μM pH 7	pH 4	pH 10
14-dC-A + 14-dC-S	57		
14-dC-A + 14-BrdC-S	63		
14-dC-S + 14-BrdC-A	62.5		
14-BrdC-A + 14-BrdC-S	65		
12-dC-A + 12-dC-S	60	53.5	50
12-dC-A + 12-BrdC-S	69.5	70	54
conclusion: $\Delta G°_d$ (kcal/mol)	0.4	0.68	0.17

(b) Kinetics: Blunt Ends

oligodeoxynucleotides	temp (°C)	half-time (min) for displacement with 12-BrdC-S at 101 μM	20 μM	4 μM
12-dC-S[a] (C = 0.25 μM) + 12-dC-A (C = 0.75 μM)	37	2	4–8	16–32
	32	4–8	8–16	32
	27	4–8	32–64	128–256

(b) Kinetics: Overhangs

oligodeoxynucleotides	temp (°C)	half-time (min) for displacement with 16-BrdC-S at 3 μM	0.57 μM
12-dC-S[a] (C = 0.25 μM) + 16-dC-A (C = 0.57 μM)	47	<1	4–8
	37	<1	16–32
	27	<1	2–4
	22	<1	<1

[a] 5′ ^{32}P labeled.

containing oligodeoxynucleotides, respectively. Let $\Delta G°_d$ be the difference in free energy between the free energy for dissociation of a BrdC·dG base pair and a dC·dG base pair. Given that $T°$ is 298.16 K and the helix–coil transition enthalpy, $\Delta H°$, is 94.8 or 98.8 kcal/mol for 12-mers or 14-mers, respectively (Quartin & Wetmur, 1989), then for a duplex containing N_{BrdC} BrdC-dG base pairs

$$\Delta G°_d = [(T°\Delta H°/T_m T_m')(\Delta T_m)]/N_{BrdC}$$

if $\Delta H°$ is constant. If, instead, $\Delta S°$ is constant, then

$$\Delta G°_d = [(\Delta H_{av}/T_{m,av})(\Delta T)]/N_{BrdC}$$

where "av" is used to designate average values. The actual $\Delta G°_d$ should be between these two extremes, which differ by only approximately 10–15%. The average value for $\Delta G°_d$ derived from the data in Table IIa is 0.4 kcal/mol at pH 7. Thus, allowing $\Delta H°$ to be constant, $\Delta G°_d = RT° \ln K = 400$ or $K = 1.9$, where K is the relative stability of a BrdC·dG base pair compared to a dC·dG base pair. The value of $\Delta G°_d$ (and hence K) is pH-dependent because of the difference between the acid–base dissociation constants of BrdC and dC.

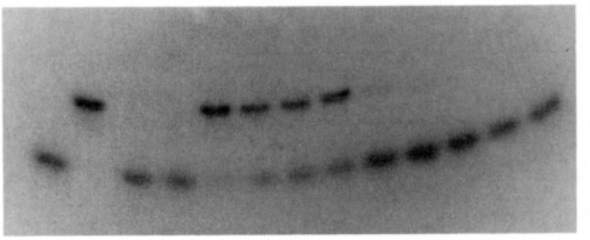

FIGURE 2: Displacement of a dC oligodeoxynucleotide by its BrdC analogue in 1 M NaCl. Lane 1: 5′-^{32}P-labeled 12-dC-S only. Lane 2: With 12-dC-A. Lanes 3 and 4: 12-dC-S/12-dC-A duplex incubated at 32 °C with 400 μg/mL 12-BrdC-S for 128 min (lane 3) and 256 min (lane 4). Lanes 5–13: The same duplex incubated with 80 μg/mL 12-BrdC-S for 1, 2, 4, 8, 16, 32, 64, 128, and 256 min, respectively.

BrdC-containing oligodeoxynucleotides were examined for their ability to displace their dC-containing analogues from duplexes. The oligodeoxynucleotides used in these studies appear in the second and third sections of Table I. Figure 2, lanes 5–13, shows the time dependence of the displacement of 5′-^{32}P-labeled 12-dC-S from the blunt-ended 12-dC-S·12-dC-A duplex by 80 μg/mL 12-BrdC-S. Lanes 1 and 2 show 12-dC-S alone and in a duplex (reduced mobility), respectively. Table IIb presents a summary of displacement rates. The reactions initiated at blunt ends proceed faster as the temperature increases due to increased breathing at the blunt ends. Displacement reactions where initiation occurred at four-base overhangs were more than 2 orders of magnitude faster than reactions initiated at blunt ends. The stability of a duplex formed between the displacing strand and the overhang of a preformed duplex increases with decreasing temperature. Thus, in the temperature range of 22–37 °C, the rate of displacement decreases with increasing temperature. Apparently, at 47 °C the displacement reaction proceeds by a different mechanism, perhaps involving the dissociation of the preformed duplex due to the proximity of its melting temperature.

Branch Migration Mediated Capture of Linker Oligodeoxynucleotides. The oligodeoxynucleotides listed in the

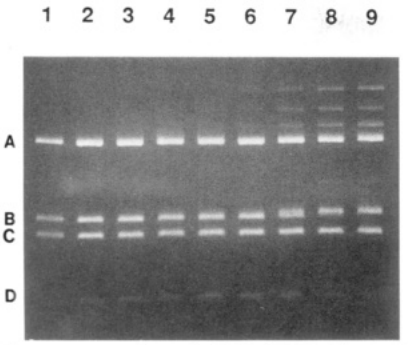

FIGURE 3: Capture reaction of 34-BrdC-D plus 14-dC-L. UV fluorogram of 1% agarose gel. Lane 1: *Rsa*I/*Pst*I-digested pALA-D (200 ng). A–D refer to fragments shown in Figure 1. Lanes 2–9: Products following ligation in the presence of 34-BrdC-D (6 μg/mL), 14-dC-L (2 μg/mL), and 5 units/mL ligase for 1, 2, 4, 8, 16, 32, 64, and 128 min, respectively.

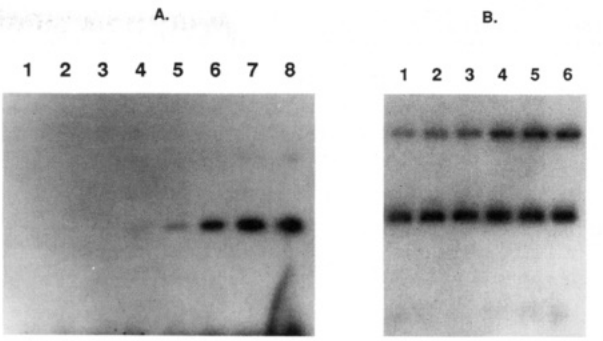

FIGURE 4: (A) Autoradiogram of Figure 3. Lanes 1–8 correspond to the radiolabeled lanes 2–9 of Figure 3. (B) Autoradiogram of capture reaction with 500 units/mL ligase for 1, 2, 4, 8, 16, and 128 min, respectively.

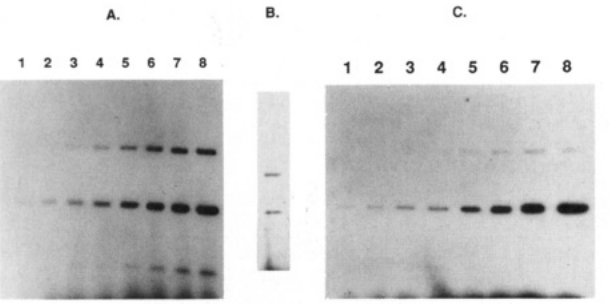

FIGURE 5: (A) Autoradiogram using same conditions as Figure 3, except that the displacer is 34–dC-D. (B) Autoradiogram at early time point using 34-BrdC-D-E(10). (C) Autoradiogram using the same conditions as panel A, except that the displacer is 34-BrdC-D-E(24).

fourth section of Table I were used for branch migration mediated capture. Plasmid pALA-D, when digested with *Rsa*I and *Pst*I, yields seven fragments, four of which contain one *Pst*I end with a 3′-overhang and one blunt *Rsa*I end (Figure 1A). Figure 1B depicts the basic experiment. The standard displacer sequence, 34-dC-D, is complementary to the first 24 nucleotides of the 3′-overhanging strand of the *Pst*I end of restriction fragment B. The linker sequence, 5′-^{32}P-labeled 14-dC-L, forms a duplex with the 5′ portion of the displacer sequence which lacks homology to the recipient duplex. Initiation of displacer–linker duplex attachment to the recipient duplex begins with base pairing of the displacer with the four-base overhang of the *Pst*I site. The displacer may replace any or all of the homologous portion of the recipient duplex strand by single-stranded DNA branch migration. DNA ligase is now used to covalently attach the kinased linker oligonucleotide to the recipient duplex, resulting in capture of the transient branch migration intermediate.

Figure 1C shows the potential for branch migration of the displacer into pALA-D fragments with one *Pst*I and one blunt (*Rsa*I) end. For fragment B, where there is full complementarity to the displacer, the maximum number of displacer–recipient base pairs (m) is 24. The maximum number of branch migration steps is ($m - 4$) because four base pairs are formed between the displacer and the 3′ *Pst*I overhang. Fragment D and the smallest (71 base pair) *Pst*I–*Rsa*I fragments contain no complementarity beyond the *Pst*I site. Thus, $m = 5$, including the four-base overhang and one branch migration step. Fragment A contains four additional nucleotides adjacent to the *Pst*I site which are complementary to the displacer. Thus, $m = 9$, including the four overhang and five branch migration steps. Unlike fragment B, where branch migration may go to completion, all branch migration structures with fragments A and D contain two single-stranded branches.

Figure 3, lane 1, shows the electrophoretic separation of the fragments of pALA-D produced by digestion with *Pst*I and *Rsa*I. Fragments A–D are indicated. Fragment C contains no *Pst*I overhang.

Figure 3, lanes 2–9, demonstrates ligation of the pALA-D fragments in the presence of the displacer 34-BrdC-D (6 μg/mL) and the linker 14-dC-L (2 μg/mL) with 5 units/mL ligase. 34-BrdC-D has the same sequence as 34-dC-D, but all the dC nucleotides have been replaced by BrdC. Ligation results in a small decrease in the mobility of fragment B. Branch migration mediated capture is quite efficient. This reaction is 50% completed in 32 min. A similar experiment carried out by using 25 units/mL ligase proceeded 50% to completion in 8 min (data not shown). Thus, the reaction rate was linearly dependent on ligase concentration.

Table III: Linker Capture

	(a) Classes I and II displacer			
	34-BrdC-D		34-dC-D	
fragment	exptl	calcd	exptl	calcd
A	16	9	4	4
B	384	344	16	15
D	1[a]	1[a]	1[a]	1[a]

	(b) Class III rel yield (fragment B/fragment A)	
displacer	exptl	calcd
34-BrdC-D	24	39
34-BrdC-D-E(10)	1	1
34-BrdC-D-E(24)	24	39

[a] By definition.

Figure 4A is an autoradiogram of Figure 3 based upon detection of ligation of 5′-^{32}P-labeled linker to the recipient fragments. Under conditions where the reaction with fragment B ($m = 24$) is carried to completion, fragment A ($m = 9$) is barely visible and fragment D ($m = 5$) cannot be detected. These results demonstrate conditions for maximum specificity of linker capture. Figure 4B presents the results of ligation in the presence of 500 units/mL ligase. Under these conditions, the capture of the linker by fragment A, although proceeding more slowly than capture by fragment B, tended toward completion, and some linker capture by fragment D was detected.

Figure 5A shows an autoradiogram analogous to Figure 4A, where the displacer strand is 34-dC-D instead of 34-BrdC-D. Under conditions where the reaction with fragment B is carried

50% to completion, on the basis of the fluorogram of the gel (data not shown), both fragment A and fragment D were easily detected. Table IIIa contains relative autoradiographic intensities for fragments A, B, and D, with 34-BrdC-D and 34-dC-D as displacer molecules. The intensity of fragment D was assigned the value of 1. Since the intensities for fragments A and B were determined by using a 2-fold time course assay, they are reliable to ±50%. The data in Table IIIa demonstrate that the use of BrdC is important for obtaining highly specific branch migration mediated capture.

Displacer 34-BrdC-D-E(10) was synthesized with one incorrect nucleotide such that only 9 base pairs could be formed between the displacer and the complementary recipient strand prior to the mismatch. Figure 5B is an autoradiogram of a branch migration mediated capture reaction carried out under conditions leading to trace levels of capture. The intensity ratios are given in Table IIIb. The capture by fragment B (m = 24 with mismatch at 10) was exactly the same as that of fragment A (m = 9). Thus, a single mismatch at position 10 blocked subsequent branch migration.

Figure 5C shows an autoradiogram analogous to Figure 4A, where the displacer was 34-BrdC-D-E(24). The final nucleotide of 34-BrdC-D-E(24) differed from that of 34-BrdC-D. The results were similar to those Figure 4A. Thus, the formation of the final singly branched structure (Figure 1B, lower right) is unnecessary to achieve the high specificity observed for capture reactions of BrdC-containing oligodeoxynucleotides.

Cloning and Sequencing. The product of a branch migration mediated capture reaction was ligated into a vector and cloned in *E. coli* in order to demonstrate that neither the presence of BrdC nor the presence of a branched structure interfered with the cloning process. The plasmid, pALA-D, was digested with *Pst*I, leading to 2771- and 1037-nucleotide fragments. The 2771-nucleotide pUC9-containing fragment, was inactivated as a cloning vector by digestion with *Sma*I, which leads to two blunt-ended fragments. Each end of the 1037-nucleotide fragment terminated in a *Pst*I site, one of which was the same as the *Pst*I site in fragment B. The branch migration mediated capture reaction was carried out as described above by using 34-BrdC-D and 14-dC-L. The products were ligated into pUC19 digested with *Sph*I and *Pst*I, both of which cut in the polylinker. The *Sph*I end contained a 3′-overhang complementary to the overhang created by the capture of the displacer–linker duplex. Four independent plasmid clones were digested by various restriction endonucleases and shown to have the patterns expected for linker incorporation. One of these clones was chosen for double-stranded DNA sequencing. The autoradiogram of the sequencing gel depicted in Figure 6 confirmed the capture reaction with incorporation of the linker sequence. These results suggest that *E. coli* contains polymerases and nucleases capable of handling both BrdC substitutions and branched structures.

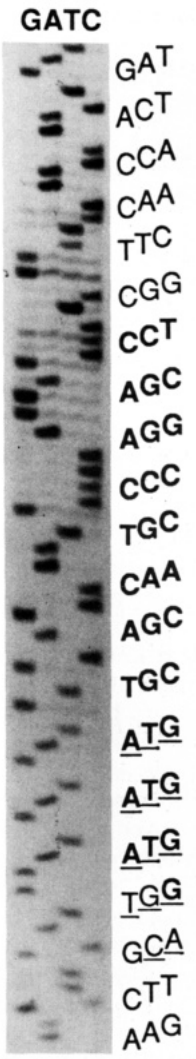

FIGURE 6: Autoradiogram of sequencing gel showing the region of incorporated displacer (bold) and linker (underlined) sequences.

THEORY

The theoretical analysis of branch migration mediated capture reactions allows generalization of the results presented above to any set of linker and displacer molecules. In our experiments, the branch migration event was initiated by formation of four base pairs with a 3′-overhang remaining after cleavage of DNA with *Pst*I. These base pairs are numbered 1–4. Additional base pairs would form only at the expense of existing base pairs, a branch migration event. For the purposes of this calculation, assume that the free energy of any structure involving two single-stranded branches is independent of the location of the branch point. Three calculations are presented below, based on principles of statistical mechanics of nucleic acids outlined in Cantor and Schimmel (1980). Class I reactions occur when the displacer strand is homologous to the displaced strand, permitting branch migration to proceed to completion. Branch migration mediated capture of the displacer (34-BrdC-D or 34-dC-D)–linker duplex by pALA-D fragment B is a class I reaction. Class II reactions occur when branch migration is limited to a restricted number of base pairs beyond the *Pst*I site overhang. Branch migration mediated capture of the same displacer–linker complexes by pALA-D fragments A and D is an example of a class II reaction. Class III reactions occur when a nonidentical (or nonequivalent) base is introduced into the displacer strand. Branch migration mediated capture of a mismatched displacer [34-BrdC-D-E(10)]–linker complex by pALA-D fragment B is a class III reaction.

Let the following definitions apply: q = partition function; σ = initiation parameter (includes base stacking at a single strand break); C = oligodeoxynucleotide concentration (molar); s_k = equilibrium constant to add base k to a stack [$s \gg 1$ (ignore fraying)]; Ω = relative stability for two versus one single-stranded branch; K = equilibrium constant for a BrdC·dG base pair versus a dC·dG base pair (this is the same K that was determined from the melting temperature analyses); B_k = 1 if branch migration results in no nucleotide substitution; $B_k = K$ if branch migration replaces dC by BrdC; and μ = relative stability of a mismatched base pair.

For class I (complete branch migration to position m):

$$q_1 = 1 + \sigma C\{\prod_{k=1}^{4} s_k(1 + \prod_{k=5}^{m} B_k + \Omega\sum_{i=5}^{m-1}\prod_{k=5}^{i} B_k)\}$$

For class II [incomplete branch migration: a *Pst*I site overhang (positions 1–4) and additional (5 through m) branch migration]:

$$q_2 = 1 + \sigma C\{\prod_{k=1}^{4} s_k(1 + \Omega\sum_{i=5}^{m}\prod_{k=5}^{i} B_k)\}$$

For class III (branch migration with a single mismatch at site j costing μ):

$$q_3 = 1 + \sigma C\{\prod_{k=1}^{4} s_k[1 + \mu\prod_{k=5}^{m} B_k + \Omega(\sum_{i=5}^{j-1}\prod_{k=5}^{i} B_k + \mu\sum_{i=j}^{m-1}\prod_{k=5}^{i} B_k)]\}$$

Assuming that the majority of all unligated sites lack bound displacer–linker duplexes, the ratio of capture (ligation) of oligodeoxynucleotides at any two sites is given by

$$r_i/r_j = (q_i - 1)/(q_j - 1)$$

We take as the standard a *Pst*I site followed by no additional similarity (e.g., fragment D of pALA-D):

$$r_j = r_2(m = 5)$$

$$r_1/r_2(m = 5) = (1 + \prod_{k=5}^{m} B_k + \Omega\sum_{i=5}^{m-1}\prod_{k=5}^{i} B_k)/(1 + \Omega B_5)$$

$$r_2(\text{all } m)/r_2(m = 5) = (1 + \Omega\sum_{i=5}^{m}\prod_{k=5}^{i} B_k)/(1 + \Omega B_5)$$

$$r_3/r_2(m = 5) = [1 + \mu\prod_{k=1}^{m} B_k + \Omega(\sum_{i=5}^{j-1}\prod_{k=5}^{i} B_k + \mu\sum_{i=j}^{m-1}\prod_{k=5}^{i} B_k)]/(1 + \Omega B_5)$$

These ratios depend only on the sequence ($B_k = K$ if BrdC or $B_k = 1$ if not) and the parameters Ω (two versus one branch) and μ (cost of a mismatch). Calculated intensity ratios in Table III were obtained by using $K = 1.7$ (at 16 °C in ligase buffer) and $\Omega = 3$. The value of 1.7 was chosen as a best fit to the 34-BrdC-D data and was consistent with the value of 1.9 found in 1 M NaCl at 60 °C. The results were insensitive to the choice of Ω as long as $\Omega \geq 2$. A value of $\Omega > 1$ implies that a structure with two single-stranded branches is more stable than a structure with one branch. The results with the displacer 34-BrdC-D-E(10) showed that a single mismatch blocked subsequent branch migration. Thus class III reactions are the same as class II reactions terminating at the mismatch (μ is very small). In summary, the specificity of all branch migration mediated capture reactions may be predicted by using $\Omega = 3$ and the sequence of the displacer.

Discussion

We have confirmed that substitution of bromodeoxycytidine for deoxycytidine increases DNA duplex stability and have derived the pH-dependent values for the average change in free energy per substitution ($\Delta G°_d$). At pH 7.0 in 1 M NaCl, $\Delta G°_d = 0.4$ kcal/mol, so the relative stability of a BrdC·dG base pair versus a dC·dG base pair (K) is 1.9. This value was in qualitative agreement with the value which fit the capture data at pH 7.0 in ligase buffer ($K = 1.7$).

We have also demonstrated that displacement reactions using BrdC-substituted oligodeoxynucleotides are rapid, especially if initiated at a dangling end. In fact, when four complementary deoxynucleotides were added to both the BrdC-containing strand and the complementary strand in the duplex, the rate constant, k_2 (units of nucleotides), was 440 M^{-1} s^{-1} at 27 °C and increased with decreasing temperature. This rate of displacement is of the same order of magnitude as DNA reassociation with nucleation limited to a four-base region (Wetmur, 1976). The increase in the rate constant with decreasing temperature is probably due to an increase in duplex stability between the region of the overhang and the BrdC-containing displacer. These thermodynamic and kinetic measurements formed the basis of a system that permits sequence-dependent capture of displacer–linker complexes.

Capture of a displacer–linker duplex at the end of a recipient duplex can be carried to 100% completion by the addition of DNA ligase to the complexes. Our capture experiments were carried out under conditions where the half-time to achieve equilibrium between a displacer–linker duplex and a recipient duplex end was about 1 min for oligonucleotide concentrations of 6 μg/mL displacer and 2 μg/mL linker. Thus, the rate-determining step was ligation. The results presented in this work are limited to recipient duplexes with 3′-overhangs. There is no reason to believe that recipient duplexes with 5′-overhangs would behave differently. The reduced rate found for displacement reactions at blunt ends with oligodeoxynucleotides indicates that, for branch migration mediated capture at a blunt end, ligation may not be the rate-determining step.

We have found that BrdC-substituted displacer molecules were required to obtain a high specificity for class I reactions compared to capture of displacer–linker duplexes by ends with limited complementarity to the displacer molecule. In addition, the introduction of a single mismatch in the displacer molecule was found to be sufficient to block branch migration. We have demonstrated that the BrdC-containing, branched structure resulting from the capture reaction can be ligated into a cloning vector and propagated in *E. coli*. For the purposes of this demonstration, no attempt was made to remove the competing displacer–linker duplexes prior to the second ligation. Thus, branch migration of a BrdC-substituted displacer oligodeoxynucleotide followed by ligation-mediated linker capture can be the basis of a selective cloning technique.

There are two general applications for our capture system. First, specific restriction fragments can be labeled to facilitate their detection following gel electrophoresis without the need for blotting and Southern hybridization. Fluorescent bases could be incorporated or the linker oligonucleotide could be end-labeled with ^{32}P using either T4 polynucleotide kinase, as described above, or terminal deoxynucleotide transferase if higher specific activity is desired (Church & Kieffer-Higgins, 1988). The ability to label one side of a unique restriction endonuclease cleavage site could find application to cosmid mapping. The linker could also contain biotinylated nucleotides, which could be used for purification of a particular fragment by affinity chromatography (Rigas et al., 1986).

The second application of the capture system is branch migration mediated DNA cloning. Linker capture at both ends of a particular fragment would increase the cloning selectivity to the product of the selectivity at each end (i.e., to 100000-fold). The high selectivity would require the design of two sets of displacer and linker molecules, one for each end of a fragment of interest. Capture at both ends is more important for cloning than for end labeling, where selectivity could only be increased to the sum of the label captured at the two ends.

The work described in this study was limited to the use of synthetic BrdC-containing oligodeoxynucleotides. Any other nucleotide that increased duplex stability and that did not

interfere with DNA cloning could substitute for BrdC. However, at the present time, BrdC is the only useful nucleotide available as a phosphoramidite. Furthermore, BrdC may be incorporated into DNA enzymatically (Kornberg, 1980). For example, enzymatically synthesized oligodeoxynucleotides could be used to simplify chromosome walking. A displacer strand could be produced by runoff replication (Saiki et al., 1988) where only the 5′ sequence prior to the 3′-overhang restriction site is known.

Our capture technology also could simplify repetitive cloning for the purpose of examining allele-specific gene polymorphisms or mutations in a diploid organism. The polymerase chain reaction (Saiki et al., 1988) is a powerful tool for obtaining relatively short sequences of genes of interest. However, the assignment of such sequences to one or the other allele is problematic. The only practical method for obtaining the contiguous sequence of a particular allele of a eukaryotic, autosomal gene (e.g., 10–50 kb) would be to obtain allele-specific clones prior to sequencing. Branch migration mediated DNA cloning may be carried out with cloning vectors that accept either small or large fragments of chromosomal DNAs, including plasmids, bacteriophages, cosmids, and yeast artificial chromosome vectors.

References

Beattie, K. L., Weigand, R. C., & Radding, C. M. (1977) *J. Mol. Biol. 116*, 783–803.

Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry*, Chapter 23, Freeman, San Fransisco.

Church, G. M., & Kieffer-Higgins, S. (1988) *Science 240*, 185–188.

Green, C., & Tibbetts, C. (1981) *Nucleic Acids Res. 9*, 1905–1918.

Holloman, W. K., Wiegand, R., Hoessli, C., & Radding, C. M. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2394–2398.

Inman, R. B. (1964) *J. Mol. Biol. 9*, 624–637.

Inman, R. B., & Baldwin, R. L. (1964) *J. Mol. Biol. 8*, 452–469.

Kornberg, A. (1980) *DNA Replication*, p 119, W. H. Freeman, San Francisco.

Lee, C. S., Davis, R. W., & Davidson, N. (1970) *J. Mol. Biol. 48*, 1–22.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers 26*, 1601–1620.

Michelson, A. M., Massoulie, J., & Guschlbauer, W. (1967), *Prog. Nucleic Acid Res. Mol. Biol. 6*, 84–141.

Quartin, R. S., & Wetmur, J. G. (1989) *Biochemistry 28*, 1040–1047.

Radding, C. M., Josse, J., & Kornberg, A. K. (1962) *J. Biol. Chem. 237*, 2869–2876.

Radding, C. M., Beattie, K. L., Holloman, W. K., & Weigand, R. C. (1977) *J. Mol. Biol. 116*, 825–839.

Rigas, B., Welcher, A. A., Ward, D. C., & Weissman, S. M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9591–9595.

Robberson, D. L., Kasamatsu, H., & Vinograd, J. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 737–741.

Saiki, R. K., Gelfand, D. H., Stoffel, B., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., & Erlich, H. A. (1988) *Science 239*, 487–491.

Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.

Thomas, M., White, R. L., & Davis, R. W. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2294–2298.

Vary, C. P. (1987) *Nucleic Acids Res. 15*, 6883–6897.

Wetmur, J. G. (1976) *Annu. Rev. Biophys. Bioeng. 5*, 337–361.

Wetmur, J. G., Bishop, D. F., Cantelmo, C., & Desnick, R. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7703–7707.