© 2001 Oxford University Press Nucleic Acids Research, 2001, Vol. 29, No. 2 e6

Efficient construction of long DNA duplexes with internal non-Watson–Crick motifs and modifications

Xiaofeng Zheng and Philip C. Bevilacqua*

Department of Chemistry, The Pennsylvania State University, University Park, PA 16802, USA

Received September 14, 2000; Revised and Accepted November 11, 2000

ABSTRACT

We have developed a semi-synthetic approach for preparing long stretches of DNA (>100 bp) containing internal chemical modifications and/or non-Watson–Crick structural motifs which relies on splint-free, cell-free DNA ligations and recycling of side-products by non-PCR thermal cycling. A double-stranded DNA PCR fragment containing a polylinker in its middle is digested with two restriction enzymes and a small insert (~20 bp) containing the modification or non-Watson–Crick motif of interest is introduced into the middle. Incorrect products are recycled to starting materials by digestion with appropriate restriction enzymes, while the correct product is resistant to digestion since it does not contain these restriction sites. This semi-synthetic approach offers several advantages over DNA splint-mediated ligations, including fewer steps, substantially higher yields (~60% overall yield) and ease of use. This method has numerous potential applications, including the introduction of modifications such as fluorophores and cross-linking agents into DNA, controlling the shape of DNA on a large scale and the study of non-sequence-specific nucleic acid–protein interactions.

INTRODUCTION

Long stretches of double-stranded (ds)DNA (>100 bp) have uses in many applications, including studies of DNA bending, protein–DNA interactions and the construction of materials for nanotechnology (1,2). Manipulation of DNA functional groups and secondary structure offers the potential to control and study these processes in detail. Unfortunately, the preparation of long stretches of duplex DNA containing internal chemical modifications, such as ribose sugars, and non-Watson–Crick motifs, such as bulges and internal loops, is complicated by several factors. Recombinant DNA methodologies relying on cloning vectors using *Escherichia coli*, for example, cannot be used to prepare such molecules since modifications will be lost and unpaired nucleotides will either be lost or repaired. Ligation of restriction fragments without cloning into a host organism is plagued by the occurrence of numerous side-products. For example, since ligation reactions are typically intermolecular in these instances, identical fragments can oligomerize via identical restriction sites. In addition, these side-products cannot be avoided by phosphatase treatment of the recipient cloning sites or the insert since there are no host enzymes to repair these sites.

Other methods of DNA preparation rely on polymerase-based approaches such as PCR, DNA splint-mediated ligations or direct chemical synthesis (3). However, these methods have shortcomings as well. PCR-only methods preclude the straightforward introduction into long DNA of internal modifications, chimeric stretches and non-Watson–Crick motifs. DNA splint-mediated ligations are useful for introducing modifications near the end of a strand when only one ligation is required, but internal modifications typically require two or more ligations and are often inefficient, with overall yields as low as a few percent (4,5). Direct chemical synthesis of the individual strands is generally limited to strand lengths under ~100 nt. Moreover, formation of duplexes from the component strands requires annealing of long strands, which can lead to numerous side-products caused by misannealing and self-structure of the individual strands.

As an alternative, we designed a semi-synthetic method in which the individual strands are not handled. Rather, the component strands were prepared as duplexes and side-products were recycled by one-pot ligation/restriction digestions. To demonstrate this method, we constructed 232 bp duplexes with internal inserts of dsDNA, bulged dsDNA and ribose-containing dsDNA. The techniques described are simple to use and allow preparation of long stretches of DNA with internal chemical and secondary structure modifications. These molecules should find use in a wide variety of biochemical studies.

MATERIALS AND METHODS

Construction of chimeric nucleic acids

All-DNA oligonucleotides were from Integrated DNA Technologies unless otherwise noted; the ribose-containing top strand DNA was from Dharmacon Research; restriction enzymes were from New England Biolabs.

PCR was carried out for 20 cycles (95°C for 1 min, 54°C for 1 min, 72°C for 1 min) using pUC19 as the template to provide a 229 bp dsDNA fragment with the polylinker in the middle. Primers used in PCR were: TS229, 5′-GCTATTACG-CCAGCTGGCGAAA; BS229, 5′-GCTCGTATGTTGTGT-GGAATTGTGAG. PCR products were run on a 2% agarose gel (FMC BioProducts) next to a molecular weight ladder and gel purified without denaturation (Zymoclean). The PCR

*To whom correspondence should be addressed. Tel: +1 814 863 3812; Fax: +1 814 863 8403; Email: pcb@chem.psu.edu

fragment was double digested with *Eco*RI and *Bam*HI to release insert 1 (17 bp), and the 5′- (101 bp) and 3′-fragments (103 bp) were co-purified. (The number of base pairs given does not count the 5′-single-stranded sticky end.)

Inserts were prepared by annealing equimolar amounts of short complementary DNA strands at 95°C for 2 min in 10 mM Tris, pH 7.5, 1 mM EDTA, 150 mM NaCl, which resulted in the appropriate sticky ends. The bottom strand for all three inserts was 5′-pGATCATGAGACCGTGGCCAACCCC, where p denotes a 5′-phosphate. The top strands were as follows: 5′-pAATTGGGGTTGGCCACGGTCTCAT for the dsDNA insert; an extra A was inserted between the GG and TC for the non-Watson–Crick bulge motif; the T of GGTC was replaced with a rU for the internal chemical modification. Concentration of the double-stranded insert was 0.05 µM and concentrations of the 5′- and 3′-fragments were 0.025 µM each, determined spectrophotometrically. This was found to be the optimal ratio of insert to fragments. Inserts were ligated into the 5′- and 3′-fragments with T4 DNA ligase and side-products were recycled by including all of the following restriction enzymes in one reaction: *Bam*HI, *Eco*RI, *Mfe*I and *Bcl*I. Enzymes were as follows (in a total reaction volume of 30 µl): 1.5 µl of 3 U/µl T4 DNA ligase (Promega); 1 µl each of 20 U/µl *Bam*HI and *Eco*RI; 1.5 µl each of 10 U/µl *Mfe*I and *Bcl*I. The buffer was Buffer 4 from New England Biolabs, which was supplemented with ATP to give final reaction conditions of 20 mM Tris–acetate, 10 mM magnesium acetate, 50 mM potassium acetate, 1 mM dithiothreitol and 1 mM ATP, pH 7.9 at 25°C. The temperature of the reaction was cycled between 16 and 37°C in a Perkin Elmer GeneAmp 2400 thermal cycler. Ligation products (232 bp) were extracted with phenol/chloroform, precipitated, triple-digested with *Eco*RI, *Bam*HI and *Mfe*I at 37°C in Buffer 4 without ATP to remove any residual side-products and purified on an agarose gel. Product duplex was either characterized by restriction digestion and sequencing or was radiolabeled at the 5′-ends (p*) with [γ-^{32}P]ATP and polynucleotide kinase (New England Biolabs) and purified by PAGE.

Characterization of chimeric nucleic acids

Inserts contained a unique *Msc*I site and formation of the correct product was verified by *Msc*I digestion. In addition, blunt-end cloning of the 232 bp ligation product was performed using the PCR-Script Kit (Stratagene). White colonies were picked and sequenced using the primer 5′-ACTAAAG-GGAACAAAAGCTG. This primer is complementary to the pPCR-Script cloning vector but does not occur in pUC19, thus preventing potential annealing within the insert itself.

The lengths of the strands from the chimeric DNA–ribose duplex and the bulged DNA duplex were examined by 10% PAGE (acrylamide:bisacrylamide ratio 29:1) in the presence of 7 M urea with 1× TBE using 5′-end-labeled samples. The shape of the bulged DNA duplex was compared to the unbulged DNA duplex by native 10% PAGE (acrylamide:bisacrylamide ratio 79:1) with 0.5× TBE at 22°C. The presence of the single ribose sugar in the chimeric duplex was tested by treating the samples with alkali (100 mM Na_2CO_3/$NaHCO_3$, 2 mM EDTA, pH 11) at 100°C for 30 min. Samples were fractionated by 6% PAGE in the presence of 7 M urea and a ribose sugar was confirmed by the presence of a cleavage product of the appropriate length.

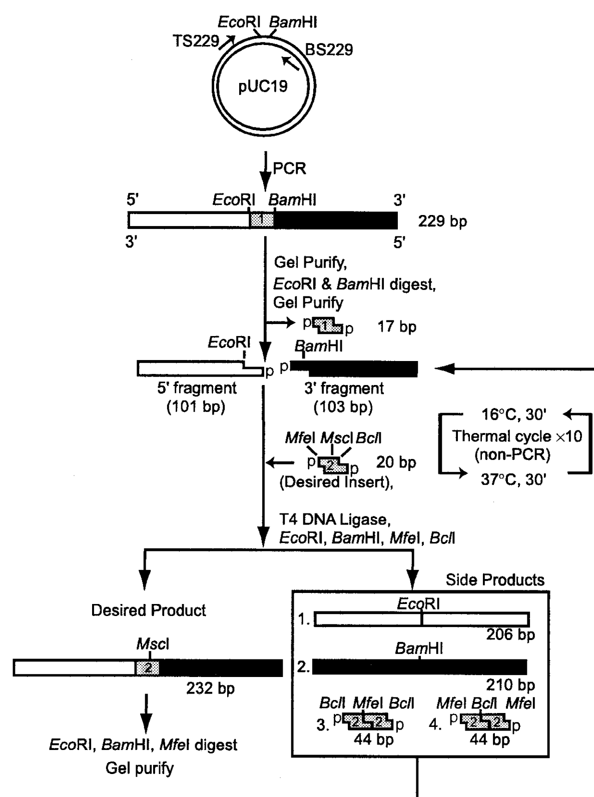

Figure 1. Strategy used to construct DNA molecules. A 229 bp PCR fragment that contains *Eco*RI and *Bam*HI cleavage sites in the middle was generated from pUC19 by 20 rounds of PCR. Primer sequences are given in Materials and Methods. After PCR the sample was purified on an agarose gel (Zymoclean) and digested with *Eco*RI and *Bam*HI to release a 17 bp fragment (insert 1, gray). The 5′-fragment (101 bp) containing an *Eco*RI sticky end (white) and 3′-fragment (103 bp) containing a *Bam*HI sticky end (black) were gel purified. The desired insert (insert 2, gray) with a *Mfe*I site on one end and a *Bcl*I site on the other was inserted between the 5′- and 3′-fragments using T4 DNA ligase. Inserts used in these experiments contained unmodified dsDNA, single bulged A-containing dsDNA or ribose-containing dsDNA. All inserts were 20 bp and had four 5′-dangling nucleotides on each end with a 5′-phosphate, denoted with a p. Since the desired product with a single insert has lost all four restriction sites, it can be enriched from the four side-products by digestion with the four restriction enzymes, followed by agarose gel purification (Zymoclean). Ten rounds of thermal cycling were used to recycle the side-products and increase the overall yield.

RESULTS AND DISCUSSION

The strategy for construction of the modified DNA is summarized in Figure 1. The first step involves preparation of a double-stranded fragment containing a polylinker in the middle. This was accomplished by PCR using pUC19 as the template and primers that flank the polylinker. In these experiments a 229 bp fragment was prepared, although the length could potentially be much greater. Next, two restriction digestions were carried out in the polylinker region to open a cloning site for the insert of interest (insert 2, Fig. 1). Insert 2 was prepared by annealing oligomers containing the modification or non-Watson–Crick motif. This annealing reaction is very reliable since it involves only short oligomers (24mer or 25mer). The insert has sticky ends for *Mfe*I and *Bcl*I, which are complementary to the *Eco*RI and *Bam*HI sticky ends, respectively. The desired product does not have restriction sites for any of these four enzymes,

whereas all possible side-products have one or more of these sites. Therefore, the entire ligation reaction mixture was simultaneously treated with *Eco*RI, *Bam*HI, *Mfe*I and *Bcl*I. The temperature of the reaction was cycled between 16 (30 min) and 37°C (30 min) a total of 10 times, since the ligase reaction is optimal at low temperatures where the DNA ends are not frayed and the restriction digestions are optimal at higher temperatures (3,6). A buffer compatible with all five enzymes was found (Materials and Methods). The insert also contains a unique restriction site for *Msc*I that is present only in the 232 bp product. The chemical and non-Watson–Crick modifications were placed off-center in the insert to maintain the *Msc*I site.

The effect of temperature cycling on the reaction yield was visualized on an agarose gel and revealed that approximately twice as much of the expected product formed with cycling than without (Fig. 2A). The effectiveness of the temperature cycling was also apparent in that the sample that was not cycled was digested to a large extent by the final triple digestion with *Eco*RI, *Bam*HI and *Mfe*I, whereas the sample that was cycled was not significantly affected by the final triple digestion (data not shown). The yields of the reactions with the three different inserts were similar at 60 ± 10% and treatment with *Msc*I resulted in complete digestion of the 232 bp DNAs to give products of the expected length (Fig. 2B). The sequence of the product containing the unmodified dsDNA insert was confirmed by blunt-end cloning and dideoxy sequencing (Fig. 2C). In addition, the three products with different inserts had identical electrophoretic mobility by denaturing gel electrophoresis, as expected (Fig. 3). The 1 nt difference in the length of the top strand of the bulged duplex was not observed, most likely due to the resolution of the gel and the presence of a bottom strand of different sequence.

To demonstrate the utility of this method we introduced a bulge or a ribose sugar into the duplex. The bulged duplex has the same electrophoretic mobility on a denaturing gel as the unbulged duplex, but significantly slower mobility on a native gel (Fig. 3A). The denaturing gel provides the mobility of a mixture of denatured radiolabeled top and bottom strands, whereas the native gel provides the mobility of the duplex. Slower mobility of the bulged duplex is consistent with the known ability of bulges to bend the helical axis of dsDNA and dsRNA (7–9).

A single ribose sugar was introduced into the top strand of the insert and analyzed by treatment with alkali (Fig. 3B). High pH conditions deprotonate the imino nitrogens and result in dissociation of the DNA strands (10) and also cause cleavage of the backbone at the position of the ribose sugar by facilitating nucleophilic attack by the 2′-hydroxyl (11). This treatment resulted in a product that migrated at the expected 120 nt position. The product was formed in 52% yield, which was close to expected since both the top and bottom strands were radiolabeled but only the top strand had a ribose sugar. Identical treatment of unmodified dsDNA resulted in no cleavage (Fig. 3B).

The method described here has numerous potential applications. It should be straightforward to prepare very long DNAs of >1000 bp containing centered or off-center non-Watson–Crick motifs such as various sized bulges and symmetrical and asymmetrical internal loops. Such molecules are good candidates for biophysical study by methods that require long

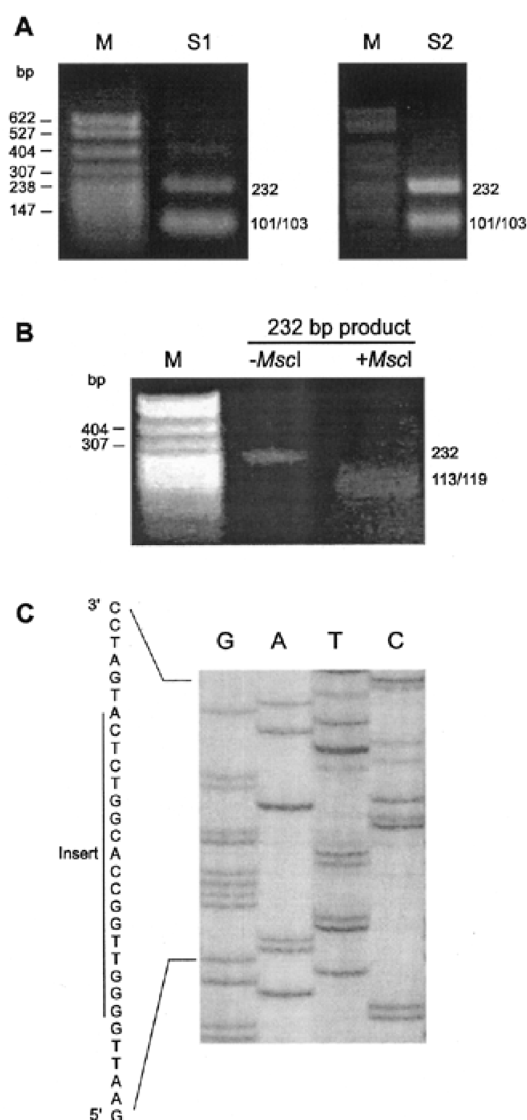

Figure 2. Analysis of reaction products. (**A**) Agarose gel analysis of reaction products. Reactions were carried out as described in Figure 1. Shown here are the results for an insert containing unmodified DNA. The sequences of the top and bottom strands of this insert are given in Materials and Methods. Similar results were found for other inserts. A 2% agarose gel is shown with ethidium bromide staining. The upper band of the samples (S1 and S2) is the 232 bp fragment (shown in Fig. 1), confirmed by *Msc*I digestion and sequencing. The lower band contains the 5′- (101 bp) and 3′-fragments (103 bp) (shown in Fig. 1) that have not been successively incorporated into the ligation. Molecular weight markers (M) from an *Msp*I digest of pBR322 DNA (New England Biolabs) are shown on the left. In sample 1 (S1) the reaction was carried out at 16°C for 16 h with no thermal cycling. In sample 2 (S2) the reaction was carried out with 10 rounds of thermal cycling between 16 (30 min) and 37°C (30 min). In both cases the sample was phenol/chloroform extracted and triple digested with *Eco*RI, *Bam*HI and *Mfe*I. (**B**) Agarose gel showing *Msc*I digestion of the 232 bp fragment. The 232 bp fragment is from S2 in (A) and the *Msc*I digestion products contain a mixture of 113 and 119 bp fragments. The 232 bp fragment is completely digested by *Msc*I. (**C**) Sequence of the ligation product and the 5′ and 3′ junctions. The DNA was sequenced by the dideoxy method. Shown is an autoradiogram of a denaturing 5% polyacrylamide gel. For the clone shown the top strand was sequenced.

substrates, including comparative electrophoresis (7), atomic force microscopy (12), electron microscopy (13) and transient

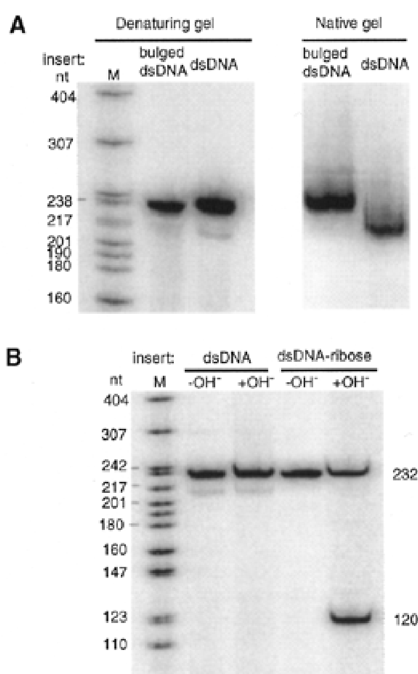

Figure 3. Characterization of the reaction products. (**A**) Comparison of electrophoretic mobilities of dsDNA and bulged dsDNA. Molecular weight marker sizes (M) are shown on the left. This marker was prepared by 5′-end-labeling an *Msp*I digest of pBR322 DNA (New England Biolabs) and boiling in formamide loading buffer prior to loading. Comparison of the electrophoretic mobilities between dsDNA and bulged dsDNA on a denaturing 10% polyacrylamide–7 M urea gel is shown on the left. The mobilities of dsDNA and bulged dsDNA show no significant differences. A 10% native gel in 0.5× TBE at 22°C is shown on the right. The mobility of bulged dsDNA is slower than dsDNA. (**B**) Hydrolysis analysis of a single ribose-containing chimeric dsDNA. dsDNA and the dsDNA–ribose were radiolabeled at the 5′-end. Samples were treated with ($+OH^-$) or without ($-OH^-$) alkali (pH 11.0) at 95°C for 30 min. Only the dsDNA–ribose sample contained a 2′-hydroxyl, as revealed by a cleavage band at the expected position.

electric birefringence (9). Since bulges induce bends in dsDNA, these substrates could be used in non-biological applications such as the design of new materials or circuits based on DNA (2). In addition, dsRNA segments could be inserted into large dsDNAs using chimeric inserts with RNA centers. This would allow biophysical study of conformational changes induced by dsRNA-binding motif-containing proteins upon binding non-sequence specifically to dsRNA (5). In the study herein a ribose sugar was introduced into the DNA to illustrate the method. Such atomic modifications potentially allow detailed molecular interactions to be probed in many different systems. Numerous other modifications could be inserted into the center of dsDNA, including primary amines and thiols, which readily couple to a wide variety of commercially available cross-linking and fluorescent molecules. It should also be possible to isolate only one of the two strands by placing a biotin at the end of the undesired strand and releasing the desired strand by denaturing chromatography using a streptavidin support (10).

ACKNOWLEDGEMENT

This work was supported by a grant from the National Institutes of Health (GM58709).

REFERENCES

1. Crothers,D.M., Drak,J., Kahn,J.D. and Levene,S.D. (1992) DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol.*, **212**, 3–29.
2. Seeman,N.C. (1999) DNA engineering and its application to nanotechnology. *Trends Biotechnol.*, **17**, 437–443.
3. Moore,M.J. and Sharp,P.A. (1992) Site-specific modification of pre-mRNA: the 2′-hydroxyl groups at the splice sites. *Science*, **256**, 992–997.
4. Strobel,S.A. and Ortoleva-Donnelly,L. (1999) A hydrogen-bonding triad stabilizes the chemical transition state of a group I ribozyme. *Chem. Biol.*, **6**, 153–165.
5. Zheng,X. and Bevilacqua,P.C. Straightening of bulged RNA by the double-stranded RNA-binding domain from the protein kinase PKR. *Proc. Natl Acad. Sci. USA*, in press.
6. Lehman,I.R. (1974) DNA ligase: structure, mechanism, and function. *Science*, **186**, 790–797.
7. Bhattacharyya,A., Murchie,A.I. and Lilley,D.M. (1990) RNA bulges and the helical periodicity of double-stranded RNA. *Nature*, **343**, 484–487.
8. Tang,R.S. and Draper,D.E. (1990) Bulge loops used to measure the helical twist of RNA in solution. *Biochemistry*, **29**, 5232–5237.
9. Zacharias,M. and Hagerman,P.J. (1995) The bend in RNA created by the trans-activation response element bulge of human immunodeficiency virus is straightened by arginine and by Tat- derived peptide. *Proc. Natl Acad. Sci. USA*, **92**, 6052–6066.
10. Bock,L.C., Griffin,L.C., Latham,J.A., Vermaas,E.H. and Toole,J.J. (1992) Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature*, **355**, 564–566.
11. Donis-Keller,H., Maxam,A.M. and Gilbert,W. (1977) Mapping adenines, guanines, and pyrimidines in RNA. *Nucleic Acids Res.*, **4**, 2527–2538.
12. Hansma,H.G., Browne,K.A., Bezanilla,M. and Bruice,T.C. (1994) Bending and straightening of DNA induced by the same ligand: characterization with the atomic force microscope. *Biochemistry*, **33**, 8436–8441.
13. Nakamura,T.M., Wang,Y.-H., Zaug,A.J., Griffith,J.D. and Cech,T.R. (1995) Relative orientation of RNA helices in a group I ribozyme determined by helix extension electron microscopy. *EMBO J.*, **14**, 4849–4859.