Enzymatic Synthesis of Region-Specific Isotope-Labeled DNA Oligomers for NMR Analysis

Georges Mer and Walter J. Chazin*

Department of Molecular Biology (MB9)
The Scripps Research Institute
10550 North Torrey Pines Road
La Jolla, California 92037

Received September 17, 1997

The recent development of an enzymatic procedure for large scale synthesis of DNA oligomers[1] uniformly labeled with the stable isotopes ^{13}C and ^{15}N will undoubtedly advance the application of NMR spectroscopy to structural studies of DNA in the same way that uniform labeling by in vitro transcription has led to the recent quantum leap in NMR structural characterization of RNA.[2−5] As first demonstrated for proteins, uniform ^{13}C and ^{15}N enrichment significantly improves spectral resolution by allowing the dispersion or filtering of ^{1}H frequencies via scalar coupling with the heteronuclei. However, most of this benefit from heteronuclear labeling is lost for nucleic acids larger than ∼30−40 residues because the limited library of nucleic acid residues results in extensive spectral overlap for all three nuclei. Additional limitations arise due to the increased relaxation rates of larger oligonucleotides, which is exacerbated by the heteronuclear dipolar couplings. These problems can be circumvented by regional labeling of the oligonucleotide, i.e., labeling of only a subset of residues within one oligonucleotide. In this communication we present a general method for region-specific labeling of DNA and its application to the study of a 64-residue DNA crossover structure that serves as a model for the Holliday junction intermediate in genetic recombination and repair.[6−8] Although the global structural features of such branched structures have been studied by a variety of techniques,[9,10] little is known about the details of DNA conformation at junctions. With specific labeling of the crossover region, this structural information can be obtained from NMR.

The application of region-specific isotopic enrichment is demonstrated by partial labeling of a 16 nucleotide oligomer (16mer) in which the 6 central residues are uniformly enriched with ^{13}C and ^{15}N, while the rest of the residues are not labeled (Figure 1). This oligomer constitutes one strand of the 4-arm model Holliday junction structure. Region-specific DNA labeling combines two polymerization reactions of deoxyribonucleoside triphosphates (dNTPs), catalyzed by a 3′-5′ exonuclease-free mutant (D424A) of the Klenow DNA polymerase. The first step utilizes an oligonucleotide hairpin structure designed to link together a five residue primer and the template needed for the enzymatic extension of the primer.[1] This first elongation reaction incorporates the six labeled nucleotides by using uniformly ^{13}C,^{15}N-enriched dNTPs. The primer corresponds to the DNA sequence at the 5′ end of the final product; it is connected to the 3′ end of the template-loop structure via a single ribonucleotide. This design enables the separation of the product (the extended primer) from the template after alkaline hydrolysis of the phosphodiester bond at 3′ of the ribonucleotide.[1] After cleavage the product is purified by anion exchange chromatography.

This half-labeled oligomer is extended with 5 nonlabeled residues by a second step of polymerization. Here we introduce the concept of using a biotinylated template, as opposed to the

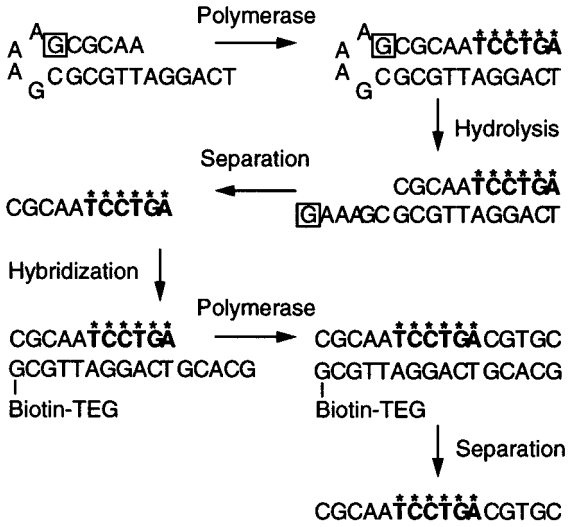

Figure 1. Two-step method for enzymatic synthesis of region-specific ^{13}C,^{15}N-enriched DNA applied to the labeling of the 6 central nucleotides of a 16-residue oligomer. The ribonucleotide residue is indicated by the boxed G. The labeled residues are highlighted with asterisks. The exonuclease-free mutant of the Klenow polymerase was expressed in *E. coli* and purified according to a published procedure.[17] The oligonucleotide templates (including those containing a single ribonucleotide residue) were synthesized on a Pharmacia LKB gene assembler with use of solid-state β-cyanoethyl phosphoramidite chemistry, purified by anion exchange chromatography with an HQ POROS resin on a Biocad Sprint (PerSeptive Biosystems), and desalted by using C18 Sep-Pak cartridges (Waters). The hairpin structure and the biotin-TEG-tagged oligomer were each synthesized using four 1 μmol scale columns (Glen Research), respectively. ^{13}C,^{15}N-labeled dNTPs were prepared from the corresponding labeled dNMPs (Martek Biosciences Corporation) by phosphorylation with a mixture of enzymes and the kinase activity from JM101 *E. coli* cells according to the procedure of Hurlbert and Furlong.[1,18] The reactions were monitored by HPLC with a Vydac nucleotide analysis column. The polymerization reactions were carried out in 10 mM Tris/HCl at pH 7.5, 0.1 mM template, 0.50 mM of each dNTP (first step reaction) and 0.75 mM (second step reaction), 5 mM $MgCl_2$, and 3 units of Klenow polymerase D424A per nanomole of DNA template. The mixtures were incubated for 3 h at 37 °C. Reactions were checked by 19% polyacrylamide gel electrophoresis. The first enzymatic reaction was performed with 0.75 μmol of hairpin DNA. The elongation reaction was followed by a desalting step with a Centriprep-3 (Amicon) centrifugal concentrator, which removes salts and unincorporated dNTPs. The product was cleaved from the template-loop in 0.4 mM KOH for 3 h at 55 °C. It was then purified by anion exchange chromatography with use of HQ POROS resin on a Biocad Sprint (PerSeptive Biosystems) in 10 mM NaOH with a gradient of 0.4−0.7 M NaCl. The purified DNA was desalted on a C18 Sep-Pak cartridge (Waters) and used for the second enzymatic reaction. The product of this reaction was purified in a similar manner. From the initial 0.75 μmol of hairpin DNA, 0.4 μmol of extended primer was obtained which then yielded 0.21 μmol of the final purified product. This is enough to prepare 0.200 mL of a 1 mM sample, sufficient for an NMR microtube (Shigemi Inc.). The overall yield of the procedure is about 28%, and it can be readily scaled to lower or higher quantity preparations.

* To whom correspondence should be addressed.

(1) Zimmer, D. P.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3091−3095.
(2) Batey, R. T.; Inada, M.; Kujawinski, E.; Puglisi, J. D.; Williamson, J. R. *Nucleic Acids Res.* **1992**, *20*, 4515−4523.
(3) Nikonowicz, E. P.; Sirr, A.; Legault, P.; Jucker, F. M.; Baer, L.; Pardi, A. *Nucleic Acids Res.* **1992**, *20*, 4507−4513.
(4) Nikonowicz, E. P.; Pardi, A. *Nature* **1992**, *355*, 184−186.
(5) Ramos, A.; Gubser, C. C.; Varani, G. *Curr. Opin. Struct. Biol.* **1997**, *7*, 317−323.
(6) Stahl, F. W. *Genetics* **1994**, *238*, 241−246.
(7) Whithouse, H. L. K. *Genetic recombination: understanding the mechanism*; John Wiley: New York, 1982.
(8) Holliday, R. *Genet. Res.* **1964**, *5*, 282−304.
(9) Seeman, N. C.; Kallenbach, N. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 53−86.
(10) Lilley, D. M.; Clegg, R. M. *Q. Rev. Biophys.* **1993**, *26*, 131−175.

S0002-7863(97)03267-8 CCC: $15.00
Published on Web 01/13/1998

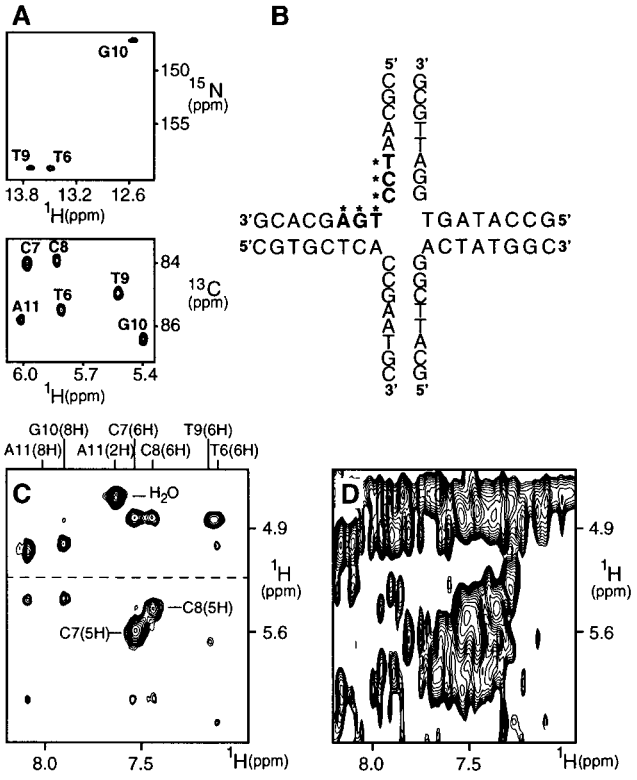

Figure 2. Heteronuclear NMR of the region-specific labeled duplex, and the nucleotide sequence and NOESY spectra of the Holliday junction model. (A) Contour plots of the imino and 1′H proton regions of the 16mer duplex are shown in the $^1H/^{15}N$ (upper panel) and $^1H/^{13}C$ (lower panel) HSQC spectra, respectively. The spectra were acquired from 0.3 mM samples in H_2O ($^1H/^{15}N$ correlation) and D_2O ($^1H/^{13}C$ correlation) at 300 K and pH 7.5 with use of a Bruker AMX500 spectrometer. (B) The nucleotide sequence of the 32-base-pair model Holliday junction J1. The labeled residues are indicated with asterisks. (C and D) The 1′H, 3′H to 6H, 8H regions of the ^{13}C NOESY−HSQC of labeled DNA (left) and 2D NOESY of unlabeled J1 (right) are shown. The dashed line separates the 1′H and 3′H regions. Experiments were acquired in H_2O solutions at 310 K and pH 7.5 with a 200 ms mixing time with use of a Bruker AMX500 spectrometer. The DNA concentrations were 0.7 and 0.5 mM for the labeled and nonlabeled samples, respectively. Water suppression was achieved in all experiments with the WATERGATE pulse scheme.[19]

template-primer hairpin approach, to solve the problem of separating the fully extended primer from the template. In this reaction, the template is a 16 nucleotide oligomer tagged at the 5′ end with biotin-triethyleneglycol (biotin-TEG). The tag increases the retention time of this template relative to the newly transcribed strand, allowing facile separation and purification of these components by anion exchange chromatography even though they are DNA oligomers of the same length. Such tagged templates can be used in any number of additional extensions because they can be recovered with high efficiency.

Both the intermediate 11mer and the final product 16mer were shown to have the correct nucleotide composition by NMR spectroscopy. The final product was first paired to a complementary nonlabeled oligomer to form a duplex. The heteronuclear spectra from this duplex shown in Figure 2A confirm the correct identity of the 16mer product $^{13}C,^{15}N$-labeled from T6 to A11 ($dC_1GCAA_5TCCTG_{10}AGCAC_{15}G$). A second sample was then prepared, wherein the oligomer was incorporated as part of the model Holliday junction J1 (Figure 2B). The ^{13}C and ^{15}N resonances were assigned on the basis of correlations with the previously assigned[11,12] protons by using heteronuclear single quantum coherence (HSQC) experiments.[13] Assignments were confirmed by comparison to three additional duplex and J1 samples prepared with oligomers containing single labels at T6 ($^{13}C1'$), T6 ($^{15}N3$), or T9 ($^{15}N3$) that were produced with standard chemical synthesis.

The Holliday junction is a large molecule for NMR studies and represents a typical system for which the application of regional labeling will be highly beneficial. Small 32 base pair analogues of the Holliday junction have been extensively analyzed by ^1H NMR, but the ability to determine their three-dimensional solution structure is severely limited by the very high degree of cross peak overlap in 2D and even 3D ^1H NOESY spectra.[11,14] Figure 2C shows a small region of a ^{13}C-edited NOESY experiment (NOESY−HSQC)[15] along with the same region from a standard 2D NOESY[16] spectrum (Figure 2D). This illustrates the huge increase in cross peak resolution made possible by region-specific labeling of the DNA. It also serves to demonstrate the advantages of region-specific labeling over other currently available methods. Chemical synthesis offers one alternative, but would be both costly and time-consuming for the preparation of either the six single labeled oligonucleotides or the one oligonucleotide with six labeled residues. The advantage of region-specific labeling over labeling the entire oligonucleotide by using the original polymerase based method is the substantial spectral simplification that results from observing signals only from the region of interest. In the case of Holliday junction structures, since the duplex arms are known to exist in standard B-form DNA conformation, this approach enables a focus solely on the critical crossover region of the molecule, without having to assign and analyze the large number of other residues. Region-specific labeling is envisioned as a powerful approach for the study of complexes with Holliday junction-binding proteins in order to obtain insights into the recognition and resolution of these structures in recombination and repair. The ability to selectively focus studies on a specific region of DNA should prove valuable in a variety of applications ranging from DNA lesions to reasonably large protein/DNA complexes.

Acknowledgment. We thank Dr. David P. Millar for his continued interest in this project, J. C. Van der Schans for his assistance in oligonucleotide synthesis and protein purification, Dr. Jeroen Pikkemaat for helpful discussions and for purifying the nonlabeled strands of the Holliday junction, Dr. Catherine Joyce for providing the expression systems for various mutants of the Klenow DNA polymerase, Dr. Akira Ono for making the $^{13}C1'$ labeled Holliday junction, and Dr. Siobhan Miick for helpful discussions. The ^{15}N-enriched thymidine phosphoramidite was synthesized by the National Stable Isotope Resource, Los Alamos. This work was supported by a grant from the National Science Foundation (MCB 9604568) to W.J.C. G.M. acknowledges the support of postdoctoral fellowships from the French association for cancer research (ARC), the Philippe foundation, and the Human Frontier Science Program Organization.

Supporting Information Available: Experimental details (4 pages). See any current masthead page for ordering and Internet access instructions.

JA973267F

(11) Chen, S.-M.; Heffron, F.; Chazin, W. J. *Biochemistry* **1993**, *32*, 319−326.
(12) Chen, S.-M.; Chazin, W. J. *Biochemistry* **1994**, *33*, 11, 453−11, 459.
(13) Bodenhausen, G.; Ruben, D. J. *Chem. Phys. Lett.* **1980**, *69*, 185−189.
(14) Carlström, G.; Chazin, W. J. *Biochemistry* **1996**, *35*, 3534−3544.
(15) Marion, D.; Kay, L. E.; Sparks, D. A.; Torchia, D. A.; Bax, A. *J. Am. Chem. Soc.* **1989**, *111*, 1515−1517.
(16) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546−4553.
(17) Joyce, C. M.; Derbyshire, V. *Methods Enzymol.* **1995**, *262*, 3−13.
(18) Hurlbert, R. B.; Furlong, N. B. *Methods Enzymol.* **1967**, *12*, 193−202.
(19) Sklenář, V.; Piotto, M.; Leppik, R.; Saudek, V. *J. Magn. Reson.* **1993**, *102*, 241−245.