Preparation of selective and segmentally labeled single-stranded DNA for NMR by self-primed PCR and asymmetrical endonuclease double digestion

Frank H. T. Nelissen, Frederic C. Girard, Marco Tessari, Hans A. Heus and Sybren S. Wijmenga*

Department of Biophysical Chemistry, Institute for Molecules and Materials, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, the Netherlands

Received April 7, 2009; Revised and Accepted June 9, 2009

ABSTRACT

We demonstrate a new, efficient and easy-to-use method for enzymatic synthesis of (stereo-)specific and segmental $^{13}$C/$^{15}$N/$^2$H isotope-labeled single-stranded DNA in amounts sufficient for NMR, based on the highly efficient self-primed PCR. To achieve this, new approaches are introduced and combined. (i) Asymmetric endonuclease double digestion of tandem-repeated PCR product. (ii) T4 DNA ligase mediated ligation of two ssDNA segments. (iii) In vitro dNTP synthesis, consisting of in vitro rNTP synthesis followed by enzymatic stereo-selective reduction of the C2' of the rNTP, and a one-pot add-up synthesis of dTTP from dUTP. The method is demonstrated on two ssDNAs: (i) a 36-nt three-way junction, selectively $^{13}$C$_9$/15N$_3$/2H$_{1,2,3,4,5,5'}$-dC labeled and (ii) a 39-nt triple-repeat three-way junction, selectively $^{13}$C$_9$/15N$_3$/2H$_{1,2,3,4,5,5'}$-dC and $^{13}$C$_9$/15N$_2$/2H$_{1,2,3,4,5,5'}$-dT labeled in segment C20–C39. Their NMR spectra show the spectral simplification, while the stereo-selective $^2$H-labeling in the deoxyribose of the dC-residues, straightforwardly provided assignment of their C1'–H2' and C2'–H2' resonances. The labeling protocols can be extended to larger ssDNA molecules and to more than two segments.

INTRODUCTION

Isotope labeling has contributed markedly to the success of multi-dimensional hetero-nuclear NMR spectroscopy in structure and dynamics investigations of proteins (1–5), RNAs (6–28), DNAs (28–37), and their complexes (38–41). For proteins and RNAs, isotope labeling has become standard practice. In sharp contrast, for (larger) ds/ssDNAs application of isotope-labeling still lags far behind, largely due to a lack of comparably simple, flexible and cost-efficient methods for their synthesis.

Labeling of DNA by chemical synthesis is versatile and has been used to produce ds/ssDNAs, labeled uniformly (29,36) or selectively/site-specifically (32,34,37,42–44) with $^{13}$C/$^{15}$N isotopes or segmentally (30,31) with $^{13}$C/$^2$H isotopes, using tailor-made phosphoramidites. However, their synthesis requires specialized chemical expertise, is labor intensive and costly. Enzymatic synthesis has therefore been more popular in the field of DNA structural biology, and in vivo (45,46) and in vitro (45,47–57) methods have been developed.

In the in vivo methods plasmids are amplified in E. coli cells grown on labeled minimal medium (45,46). The plasmids contain multiple repeats of the desired sequence with each repeat flanked by an endonuclease-sensitive site. After enzymatic digestion single double-stranded (ds) DNA fragments are released. Although this method circumvents the requirement of isotope-labeled dNTPs, it restricts the types of labeling (see below) and is relatively inefficient.

Enzymatic in vitro synthesis of isotope-labeled ssDNA was first achieved by Zimmer and Crothers by means of a template-driven fill-in reaction (47). This method employs a DNA polymerase I fill-in reaction on the single-stranded region of a hairpin-primed DNA template, using isotope-labeled dNTPs as building blocks. A ribonucleoside is placed at the 3'-terminus of the hairpin, so that upon alkaline hydrolysis the synthesized isotope-labeled ssDNA is released. Modifications leading to improved efficiency have been reported (48,49), while Mer and Chazin (50)
described a modification that allows for segmental isotope labeling. The template-driven fill-in reaction is relatively straightforward. A drawback is the need for stoichiometric amounts of expensive modified DNA template.

A second in vitro method is Endonuclease-Sensitive Repeat Amplification (ESRA) (45,55). It is based on a self-primed PCR amplification, originally called Concatemer Chain-Reaction (51,52), in which a dsDNA template of tandem repeats is elongated using (isotope-labeled) dNTPs by means of PCR in a number of cycles of heating, annealing and extension. This amplification produces long dsDNAs consisting of multimeric repeats of the target sequence. Endonuclease-sensitive sites flank the individual dsDNA repeats (45), so that restriction enzyme digestion releases the target dsDNA. Improvements have been reported in terms of sequence flexibility (56) and length of produced target dsDNA (57). It has also been proposed to attain PCR amplification of tandem repeats in a plasmid (53,54). However, this PCR protocol is more involved and most importantly, not self-primed and thus requires stoichiometric amounts of primer. The principal advantages of the self-primed PCR amplification are its simplicity and high efficiency.

A single dsDNA template leads to copious amounts of isotope-labeled dsDNA targets, instead of one DNA target per template molecule as in the template-driven fill-in reaction. Drawbacks are that only dsDNA can be produced and that the existing protocol does not allow for segmental labeling.

Both in vitro methods need isotope-labeled dNTPs as building blocks. For RNA, isotope labeling—via in vitro synthesis—is common practice, because the required rNTP building blocks are readily available with a variety of labeling patterns; the rNTPs can either be uniformly labeled with \(^{2H/13C/15N}\) isotopes (6,7,10,11,14), when extracted from bacterial cells, or (stereo-)selectively labeled with \(^{1H/13C/15N}\) isotopes (9,12,13,15,17,27), when produced via in vitro synthesis. However, for dNTPs the situation is less favorable. The dNTP production methods involve extraction of isotope-labeled dNMPs from genomic (bacterial) DNA (14,47,48,58), on an Akta Basic equipped with a HiLoad Q Sepharose junction ssDNA (59) and a 39-nt triple-repeat three-way junction ssDNA (60,61).

### MATERIALS AND METHODS

#### Folding simulations of PCR primers and DNA segments for the (segmental) labeling of ssDNA

Predictions of the thermodynamic stability of intramolecular folds, self-complementary duplexes and the (desired) hybrids of the used DNA sequences were obtained using DINAMelt (62). In the DINAMelt simulations for the PCR sequences, the DNA concentration was 0.1 μM and salt concentrations 10 mM NaCl and 2 mM Mg\(^{2+}\), comparable to the conditions during PCR amplifications. In the DINAMelt simulations for the ligation, the DNA concentration was 10 μM and the NaCl and Mg\(^{2+}\) concentrations were each 10 mM, matching the conditions during DNA ligation.

#### Synthesis of \(^{13}C_6/^{15}N_2^{2H}_{1,2,3,4,5,6}\)dTTP and \(^{13}C_5/^{15}N_2^{2H}_{1,2,3,4,5,5}\)CTP

For calculation of yields, here and below, all concentrations of purified intermediate and final products were determined by UV absorption unless otherwise stated. Prior to dTTP and dCTP synthesis, UTP residues were in vitro synthesized from 450 μmol \(^{13}C_6/^{2H}_{1,2,3,4,5,6}\)d-glucose and 440 μmol \(^{13}C_4/^{15}N_2/^{2H}_{1,2,3,4,5,6}\)-uracil (Cambridge Isotope Laboratories, Andover, MA, USA) using enzymes of the glycolysis and pentose phosphate pathway (9,12,13,63) (see Supplementary Data for a detailed description). \(^{13}C_6/^{15}N_2/^{2H}_{1,2,3,4,5,5}\)CTP residues were synthesized in a 120 mM reaction mixture containing 0.5 mM (60 μmol) \(^{13}C_6/^{15}N_2/^{2H}_{1,2,3,4,5,5}\)-UTP, 10 mM \(^{15}NH_4\)Cl and 7.5 units CTP synthase (64) (Supplementary Data). The UTP and CTP were separated from their ATP cofactor on an Akta Basic equipped with a HiLoad Q Sepharose HP column (GE Healthcare) using a linear gradient from 50–350 mM NaCl in Milli-Q water (Millipore) at pH 9. Nucleotide containing fractions were desalted on a Sephadex G10 column (0.25 x 30 cm), lyophilized and dissolved to 100 mM in Milli-Q water.

Fifty-nine micromoles CTP and 60 μmol UTP were exchanged with \(D_2O\) by lyophilization for three times in 1 ml \(D_2O\). A 5× concentrate of reduction buffer (20 mM HEPES (pH 7.5), 30 mM DTT, 0.5 mM EDTA and 1 M NaAc) was exchanged with \(D_2O\) by vacuum concentration to dryness for three times in \(D_2O\), subsequently dissolved in \(D_2O\) and the pH adjusted to 7.5 with 10 M NaOD. \(D_2O\) exchanged CTP and UTP were added to a final concentration of 1 mM in separate reactions and 5 μl of 20 mg/ml ribonucleotide triphosphate reductase (RTPR) (65,66) (Supplementary Data) along with 10 μl of 5 mM coenzyme B12 in \(D_2O\) were added per ml of reduction buffer and incubated for 1 h at 37°C in the dark. The reduction
progress was checked by reversed phase chromatography (RPC) on a PepRPC 15 HR10/10 column (GE Healthcare) by applying a 20 ml linear gradient from 50 mM TEAA (pH 7) to 40% 50 mM TEAA (pH 7) + 30% acetonitril. The reaction mixtures were purified over a Dowex 1x2 anion exchange column (Sigma-Aldrich, Zwijndrecht, The Netherlands) using a linear gradient from 0 to 1 M NaCl. Nucleotide containing fractions were desalted by ethanol precipitation and dissolved in 120 ml Milli-Q water and stored at −20°C.

dUTP was synthesized in a one-pot add-up reaction from dUTP. All reactions were checked by RPC applying a 20 ml linear gradient from 0.2 M KH2PO4 (pH 4) to 40% 0.2 M K2HPO4 (pH 4) + 30% methanol. Twenty milliliters of 1.3 mM dUTP in 20 mM Tris–HCl (pH 8.1 at 25°C) + 5 mM MgCl2 was incubated at 85°C for 1 h in the presence of 125 ng/ml His6-tagged Pfu dUTPase. The reaction mixtures were purified over a Dowex 1x2 anion exchange column (Sigma-Aldrich, Zwijndrecht, The Netherlands) using a linear gradient from 0 to 1 M NaCl. Nucleotide containing fractions were desalted and concentrated to 5 mM in Milli-Q water and stored at −20°C.

Synthesis of 13C9/15N3/2H(1,2,3,4,5,5,5)-dC selective and segmental labeled triple-repeat junction 6 wt ssDNA

The 20 nt ssDNA segment junction 6 wt-p2 (C20-C39; Table 1) for ligation was obtained by self-primed PCR (Figure 1) as for the three-way junction ssDNA with the case specific modifications as described hereafter. The unlabeled dUTTP in the PCR mixtures was substituted by 13C9/15N3/2H(1,2,3,4,5,5)-dUTTP and the amplification primers were junction 6 wt-p2-sense and junction 6 wt-p2-asense (Table 1). Blunt end digestion of the amplified DNA was carried out with 2.5 kU of PvuII (5′-CAGCTG-3′) in buffer G+ and the second digestion with 10 kU XhoI (5′-CTCGAG-3′) in buffer R+.

The purified 20 nt junction 6 wt-p2 ssDNA segment was ligated to the 19 nt junction 6 wt-p1 ssDNA segment (Table 1) in an 11 ml preparative ligation reaction containing 7.5 μM of junction 6 wt-p2, 10 μM of junction 6 wt-p1 and 10 μM of the 28 nt DNA splint (Table 1) in T4 DNA ligase buffer. The reaction was heated 2 min at 95°C and incubated 10 min at room temperature prior to the addition of 2.8 kU T4 DNA ligase and followed by ligation at 37°C for 4 h. The 39 nt ligation product was purified from denaturing PAGE as described before and lyophilized.

NMR experiments

The three-way junction ssDNA sample was dissolved in 50 mM NaCl in D2O (pD 5.6), to 0.2 mM and the full function 6 wt ssDNA was dissolved in 10 mM sodium phosphate pH 6.7 containing 0.1 mM EDTA and 7% D2O to 0.2 mM. The samples were heated at 95°C, snap-cooled on ice-water and transferred to a Shigemi NMR tube. All NMR spectra were acquired at 25°C using a Varian 600 Inova spectrometer equipped with a shielded triple-axis gradient HCN probe. A 2D H2/C2/C1 experiment was recorded for the stereo-specific assignment of the H2 of the cytidines by observation of (C1′, H2′) correlations via the relayed H2→C2→C1′ connectivity’s. The 2D spectrum was acquired with 128 scans per increment as a data matrix of 160(t1)×614(t2) complex points, with sweep widths of 9000 (13C) and 8000 Hz (1H). A 1D imino spectrum of junction 6 wt was acquired with 256 scans, a recovery delay of 2 s and an acquisition time of 80 ms. Selective water-flip-back pulses were employed to prevent saturation of the water resonance and attenuation of the imino protons signals. A 15N,1H HSQC pulse scheme was used to acquire a 1D 15N-edited spectrum of imino protons from 15N-labeled thymidine nucleotides. A 2D 15N,1H HSQC was recorded with 64 scans per increment as a data matrix of 80(t1)×1120(t2).
complex points, with sweep widths of 1000 Hz (15N) and 16000 Hz (1H).

RESULTS

We describe and demonstrate an efficient and easy-to-use self-primed PCR method for the enzymatic synthesis of isotope-labeled dsDNA and/or ssDNA, for which we introduce and combine three new approaches.

(i) Self-primed PCR followed by asymmetric endonuclease double digestion to produce ssDNA (Figure 1). This method originates from the highly efficient and straightforward self-primed PCR-based ESRA technique described by Louis et al. (45) for the synthesis of isotope-labeled dsDNA. As in ESRA, long tandem-repeated dsDNA PCR products are synthesized by self-primed PCR in a number of heating/annealing cycles using Pfu DNA polymerase. The major modification we introduce lies in the enzymatic digestion of the PCR products, which allows for separation of the PCR products into ssDNAs. Instead of using a single blunt-end digestion, we generate asymmetrical dsDNA in two successive digestions. The double-digested products migrate distinguishably on denaturing PAGE. The desired ssDNA fragment can be used for NMR structural analysis directly after purification from gel.

(ii) Segmental labeling via DNA-splinted ligation (Table 1). Segmental labeled ssDNA can efficiently be produced by T4 DNA ligase mediated ligation of a labeled to unlabeled strand of DNA assisted by a complementary DNA splint.

(iii) In vitro synthesis of isotope-labeled dNTPs (Figure 2). In vitro production of rNTPs allows for a large variety of labeling patterns including deuteration (9,12,13,63). We demonstrate that (stereo-)selective or uniformly 2H/13C/15N-labeled dNTPs are efficiently obtained by in vitro rNTP synthesis (9,12,13,63) followed by enzymatic reduction of the C2’ hydroxyl of the produced rNTPs (65,66).

The whole method is demonstrated on the synthesis of a 36-nt three-way junction ssDNA (59) (Figure 2C, left).
Sequence names shown in bold represent synthesized labeled sequences; all others are unlabeled and obtained from Biolegio (Nijmegen, the Netherlands). Nucleotides of PCR primers shown in italics represent (a part of) the desired ssDNA sequence to be obtained. Nucleotides depicted in bold represent (a part of) endonuclease sites and underlined parts represent the desired overlap to be formed during PCR and/or ligation.

### Table 1. Overview of used and synthesized ssDNA sequences

<table>
<thead>
<tr>
<th>Sequence name</th>
<th>Nucleotide sequence</th>
<th>Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>3W-sense</td>
<td>5’CGCGTGACGGCGTGGCTGGACTGCTCTGACGTTCAACCAGCCGGGCTGCAAGGGCTT3’</td>
<td>PCR</td>
</tr>
<tr>
<td>3W-asense</td>
<td>5’CGCGTGACGGCGTGGCTGGACTGCTCTGACGTTCAACCAGCCGGGCTGCAAGGGCTT3’</td>
<td>PCR</td>
</tr>
<tr>
<td>Junction 6 wt-</td>
<td>5’CAGCTGCCCTGGCTGCTCGAGACGTGCAGCTGCCTGGCTCAGC3’</td>
<td>PCR</td>
</tr>
<tr>
<td>p2-sense</td>
<td>5’CAGCTGCCCTGGCTGCTCGAGACGTGCAGCTGCCTGGCTCAGC3’</td>
<td>Ligation</td>
</tr>
<tr>
<td>Junction 6 wt-</td>
<td>5’GCGACGTCTCGAGCGGAGCGACGCGCAAAGGGCCAGCTGCCGC3’</td>
<td>Ligation</td>
</tr>
<tr>
<td>p2-asense</td>
<td>5’GCGACGTCTCGAGCGGAGCGACGCGCAAAGGGCCAGCTGCCGC3’</td>
<td>Ligation</td>
</tr>
<tr>
<td>Junction 6 wt-</td>
<td>5’GCGGAGCGACCATCCTGGTCTGCCTCTGCAGC3’</td>
<td>NMR</td>
</tr>
<tr>
<td>p1</td>
<td>5’GCGGAGCGACCATCCTGGTCTGCCTCTGCAGC3’</td>
<td>NMR</td>
</tr>
<tr>
<td>DNA splint</td>
<td>5’CAGCCCAAGGGCGACACAAAGGTGCTGC3’</td>
<td>NMR</td>
</tr>
<tr>
<td>Three-way junction A</td>
<td>5’CGTGACCGGCTTGCGGCACCTTGCCTGCTGACG3’</td>
<td>NMR</td>
</tr>
<tr>
<td>Junction 6 wt</td>
<td>5’GCCGAGCAGACACTTCTGGTGTCCCTTGGGCTGCCGC3’</td>
<td>NMR</td>
</tr>
</tbody>
</table>

Sequence names shown in bold represent synthesized labeled sequences; all others are unlabeled and obtained from Biolegio (Nijmegen, the Netherlands). Nucleotides of PCR primers shown in italics represent (a part of) the desired ssDNA sequence to be obtained. Nucleotides depicted in bold represent (a part of) endonuclease sites and underlined parts represent the desired overlap to be formed during PCR and/or ligation.

### Figure 2. (A) In vitro (deoxy)ribonucleotide triphosphate synthesis. Labeled starting materials are shown in bold and blue. Labeled rNTPs, here as intermediate products, are circled. Labeled end-product dNTPs are shown in bold. The enzymes catalyzing different reactions are circled. See text for detailed description of the reactions. (B) Overview of the in vitro synthesized labeled dNTPs. From left to right: 13C9/15N2/2H(1,2,3,4,5)-dUTP, 13C9/15N2/2H(1,2,3,4,5)-dCTP, 13C9/15N2/2H(1,2,3,4,5)-dTTP, and 13C9/15N2/2H(1,2,3,4,5)-dTTP. The red asterisks at carbon atoms indicate 13C labels; the blue asterisks at nitrogen atoms indicate 15N labels. The circled deuterium indicates the introduced stereo-selective deuteration in the sugar moiety during rNTP reduction. (C) Predicted most stable secondary structures of the three-way junction ssDNA (left) and the triple-repeat three-way junction ssDNA (junction 6 wt, right). The arrow indicates the chosen segmental site in junction 6 wt. Blue cytidine residues are 13C9/15N2/2H(1,2,3,4,5)-labeled (see also B) and red thymidine residues are 13C9/15N2/2H(1,2,3,4,5)-labeled (see also B).

(stereo-)selectively labeled on the cytidine residues and a 39-nt triple-repeat three-way junction ssDNA (60,61) (Figure 2C, right) that is (stereo-)selectively labeled on the cytidine and thymidine residues of segment C20–C39.

### Design of PCR amplification primers and segments

The PCR primer pairs are designed to contain both target sequence and endonuclease digestion sites (Figure 1 and Table 1). It is essential that during PCR amplification the desired hybridization of primers indeed occurs and alternative stable structures do not interfere. Table 2 provides an overview of the predicted folds and stabilities of the PCR primers and their hybridization products as predicted by DINAMelt (‘Materials and Methods’ section). For all four individual primers, neither unfavorable intramolecular folds nor unfavorable self-complementary duplexes are predicted. Although, a non-amplifiable hybrid is predicted to be more stable than the desired amplifiable hybrid (ΔG values inside versus outside parentheses, Table 2), it is expected (and ultimately experimentally proven) not to block PCR, because both hybrids will occur in the annealed mixture.

For the segmental labeling of the junction 6 wt ssDNA (Figure 2C, right), the molecule was divided into junction 6 wt-p1 (19-nt, residue 1–19) and junction 6 wt-p2 (20nt; residue 20–29), so that the latter encompasses the two thymidine residues present in the structurally interesting junction region. DINAMelt simulations were performed to assess the potential presence of alternative secondary structures that could interfere with ligation (Table 2). The desired hybridization with the DNA splint was predicted to be most stable.

### Synthesis of 13C9/15N2/2H(1,2,3,4,5)-dCTP and 13C9/15N2/2H(1,2,3,4,5)-dUTP

The in vitro synthesis of 13C9/15N2/2H(1,2,3,4,5)-UTP yielded 390µmol of UTP out of 450µmol 13C9/15N2-D-glucose and 440µmol 13C9/15N2-uracil, a yield of 89%.
The subsequent synthesis of $^{13}$C$_9$/15N$_3$/2H$_{(1',2',3',4',5',5')}$ CTP, out of 60 μmol UTP, yielded 59 μmol of CTP. The reduction of 60 μmol UTP and 59 μmol CTP by RTPR led to 58 μmol dUTP and 56 μmol of dCTP, yields of 96.5% and 95%. The dephosphorylation of dUTP proceeded for over 99%, the subsequent methylation of dUMP to dTMP for 93%, and the phosphorylation to dTTP for 98%, yielding 52 μmol of dTTP. The overall yields for the synthesis of dTTP and dCTP from glucose and uracil were 77% and 83%.

**Table 2.** Predicted hybridization and folding stabilities of ssDNA sequences

<table>
<thead>
<tr>
<th>DNA strand(s)</th>
<th>Intramolecular (ΔG)</th>
<th>Duplex (ΔG)</th>
<th>Hybrid (ΔG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3W-sense</td>
<td>−11.7</td>
<td>−18.5</td>
<td>NA</td>
</tr>
<tr>
<td>3W-asense</td>
<td>−7.2</td>
<td>−17.6</td>
<td>NA</td>
</tr>
<tr>
<td>3W-sense + 3W-asense</td>
<td>NA</td>
<td>−30.8* (−67.4)</td>
<td>NA</td>
</tr>
<tr>
<td>Junction 6 wt-p2-sense</td>
<td>−5.8</td>
<td>−13.8</td>
<td>NA</td>
</tr>
<tr>
<td>Junction 6 wt-p2-asense</td>
<td>−6.3</td>
<td>−17.4</td>
<td>NA</td>
</tr>
<tr>
<td>Junction 6 wt-p2-sense + Junction 6 wt-p2-asense</td>
<td>NA</td>
<td>−29.1* (−47.9)</td>
<td>NA</td>
</tr>
<tr>
<td>Junction 6 wt-p1</td>
<td>−3.0</td>
<td>−6.0</td>
<td>NA</td>
</tr>
<tr>
<td>Junction 6 wt-p2</td>
<td>−3.1</td>
<td>−9.1</td>
<td>NA</td>
</tr>
<tr>
<td>DNA splint</td>
<td>−7.4</td>
<td>−19.3</td>
<td>NA</td>
</tr>
<tr>
<td>DNA splint + junction</td>
<td>NA</td>
<td>−18.0</td>
<td>NA</td>
</tr>
<tr>
<td>6 wt-p1</td>
<td>−4.7</td>
<td>−6.0</td>
<td>NA</td>
</tr>
<tr>
<td>DNA splint + junction</td>
<td>NA</td>
<td>−17.8</td>
<td>NA</td>
</tr>
<tr>
<td>6 wt-p2</td>
<td>−3.0</td>
<td>−7.4</td>
<td>NA</td>
</tr>
<tr>
<td>DNA splint + junction</td>
<td>NA</td>
<td>−39.1</td>
<td>NA</td>
</tr>
</tbody>
</table>

ΔG values are in kcal/mol at 37°C. Simulation conditions are 10 mM NaCl, 2 mM Mg$^{2+}$ and 0.1 μM strand concentration for simulations with PCR primers and 10 mM NaCl, 10 mM Mg$^{2+}$ and 10 μM strand concentration for simulations concerning ligation. ΔG values with an asterisk represent the stability of the desired hybridization between the parts of the PCR primer pairs, whereas the ΔG value in parentheses represents the stability of a non-amplifiable hybrid between the primer pairs.

**Synthesis of selective and/or segmentally labeled ssDNA**

In step 1 of the self-primed PCR protocol (Materials and Methods section and Figure 1), PCRs with the primer pairs 3W-sense + 3W-asense and junction 6 wt-p2-sense + junction 6 wt-p2-asense, yielded primarily high molecular weight DNA (>10 kb; the main band is located near the loading slot in Figure 3A and B, lanes 1). The step 1 PCR-product of the junction 6 wt DNA (Figure 3B, lane 1) was substantially more smeared out than the step 1 PCR-product of the three-way junction DNA (Figure 3A, lane 1). The subsequent PCRs in step 2 yielded for both constructs even more high molecular weight DNAs, as evident from the reduced smearing (Figure 3A and B, lanes 1 and 2); note again that the main band of PCR products is located near the loading slot, indicating high molecular weight DNAs (Figure 3A and B, lanes 2).

This main band disappeared completely after blunt-end-generating endonuclease digestion and single bands below the 100 bp marker were observed (Figure 3A and B, lanes 3). The latter could be attributed to the expected 50 bp (3.75 mg; 0.12 μmol) and 35 bp (2.5 mg; 0.11 μmol) digestion products for the three-way junction and junction...
yield was 85%). After purification, 750 μg (0.062 μmol) of segmental selective labeled junction 6 wt ssDNA was obtained.

**NMR experiments**

Thanks to the stereo-selective deuteration of the H2′ in the three-way junction ssDNA, the stereo-specific resonance assignment of H2′ can easily be obtained via H2′(C2′)C1′ correlation (Figure 5A) in combination with H2′/C2′ correlation (Figure 5B). The labeling of dT residues in segment C20–C39 of junction 6 wt ssDNA is illustrated in Figures 5C and D. Comparison of the 1D traces in Figure 5C shows that only labeled dT imino resonances (of A:T base pairs) remain upon 15N-filtering. In the 2D HSQC (Figure 5D) only two dT imino cross-peaks are expected, given the secondary structure and the segmental labeling (Figure 2C). We observe one strong cross peak, which can be attributed to T35 in the lower stem. However, instead of the expected single cross-peak for T32, two (weaker) cross-peaks are seen, indicating multiple conformations at the junction.

**DISCUSSION**

We have shown that self-primed PCR of DNA combined with asymmetrical double digestion of the PCR product and in vitro dNTP synthesis forms an efficient and straightforward method for obtaining NMR amounts of (stereo-)selective and/or segmental 2H/13C/15N-labeled ssDNA. A host of labeling patterns is possible, thereby making NMR structural studies on (larger) ssDNA more accessible. With this labeling, NMR spectral crowding is reduced, specific resonances can be eliminated or selected and line widths reduced by deuteration.

**Synthesis of selective and/or segmentally labeled ssDNA and dNTPs**

The amplification of the DNA primers by means of self-primed PCR followed by double digestion requires attention to their design. The sequences should not become trapped in duplexes or intramolecular structures blocking PCR. When DINAMelt predicts such alternative stable structures, the spacer fragment can be shortened, lengthened or altered. The role of the spacer is essentially only to facilitate efficient restriction endonuclease digestion in the second digestion step and is of no further interest.

Furthermore, potentially any restriction site can be chosen to flank the DNA fragment of interest. Amongst the wide variety of available enzymes, it is likely that a combination can be found that generates digested ends exactly matching the 5′-end- and 3′-end of the DNA fragment of interest. Our method is based on the difference in migration distance on denaturing PAGE of the asymmetric DNA fragments that are generated thanks to the sticky-end restriction endonuclease digestion. In both described cases, the asymmetry comprises 5′-recessing ends of 4-nt. For ssDNAs up to 50-nt, this 4-nt overhang is sufficient for separation on preparative denaturing polyacrylamide gels. Larger ssDNAs can be obtained by applying the first endonuclease digestion with a

![Figure 4](https://academic.oup.com/nar/article-abstract/37/17/e114/1077706/1077706)
3'-recessing-end generating endonuclease. With the resulting 8-nt difference, ssDNA molecules up to 75-nt can be separated.

Finally, 54% and 32% of the labeled dCTPs are incorporated into the final PCR product for the three-way junction and junction 6 wt-p2. These yields are comparable to the ~50% of Louis et al. (45). Most of the minor loss could be due to instability of dNTPs over the extended heating/cooling cycles (60–95°C) in the PCR. For instance, deamination of dCTP leading to dUTP may occur, as well as aspecific dephosphorylation to dCMP (67). Finally, the in vitro synthesis of dCTP and dTTP is not limiting as they are produced with ~80% yields from their rNTP counterparts. In our experience these yields are attained when applying standard optimization parameters for PCR, such as annealing temperature, Mg2+-concentration, etc. To improve yields (e.g. the 32%) most effective would be to increase the number of PCR cycles. In the final ssDNAs for both the three-way junction and junction 6 wt-p2 ~10% of the labeled dCTP is incorporated. Part of the loss in the last step (compare with 54% and 32%) is due to the fact that the target ssDNAs contains only 38% and 45% of the cytidines present in the complete PCR repeat (see e.g. Table 1); further losses are due to a combination of incomplete digestion and purification steps.

A high yield (~95%) was obtained in the splinted ligation by T4 DNA ligase of the two DNA segments into the junction 6 wt ssDNA. We previously observed that (DNA-)splinted ligation of RNA segments using T4 DNA ligase is in general quite robust and leads to high yields (27). Nevertheless, it is crucial for successful (DNA-)splinted ligation of two ssDNA segments to verify optimal formation of the desired hybridization product.
This information can be obtained from DINAMelt simulations. Finally, our segmental labeling method can be extended to more than two segments without significant loss in ligation yield, e.g. by designing two nicks on a hybrid of three segments with a DNA splint or by using two DNA splints and three segments in one ligation reaction.

CONCLUSION

We propose and show that self-primed PCR of DNA combined with asymmetrical double digestion of the PCR product and in vitro dNTP synthesis forms an efficient and straightforward method for obtaining NMR amounts of (stereo-)selective and/or segmental $^2$H/$^{13}$C/$^{15}$N-labeled ssDNA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Funding within the context of the 6th framework program of the EU, 6th framework STREP project FSG-V-RNA is gratefully acknowledged.

FUNDING

Funding for open access charge: Department of Biophysical Chemistry of the Radboud University Nijmegen, The Netherlands.

REFERENCES

labeled DNA duplex \[\text{d}(C\text{G}A\text{G}T\text{A}\text{A}\text{G}T\text{C}^\text{13}O\text{G})_2\text{G}^\text{3}].\] Magn. Reson. Chem., 37, 203–213.


