

Structure-independent and quantitative ligation of single-stranded DNA

Tian W. Li, Kevin M. Weeks*

Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599, USA

Received 13 September 2005

Available online 18 November 2005

Abstract

Ligation of an adapter oligonucleotide to a single-stranded cDNA is central to many molecular biology techniques. Current single-stranded ligation approaches suffer from low efficiencies and are strongly inhibited by preexisting DNA secondary structure. We develop an approach for ligating low concentrations of single-stranded DNAs to a DNA adapter with near-quantitative efficiency, unaffected by secondary structure in the target DNA. This efficient DNA ligation reaction will facilitate development of robust procedures for quantifying small amounts of highly structured cDNAs and their RNA templates.

© 2005 Elsevier Inc. All rights reserved.

Keywords: cDNA; Single-stranded DNA ligation

Many molecular biology manipulations, focused especially on analyzing RNA sequences and structures in cells, require that an adapter oligonucleotide be ligated onto either the 5' or 3' end of a target cDNA derived from an RNA of interest. Important methods in this class include mapping the 5' or 3' terminus of an RNA [1,2] and performing *in vivo* analysis of RNA structure [3,4]. A serious shortcoming of current adapter ligation reactions is that the efficiency for direct ligation of an oligonucleotide to a single-stranded cDNA is variable and is often quite low. Ligation efficiencies also decrease significantly with increasing target DNA length [5]. Thus, useful but complex approaches, involving nucleotide tailing steps prior to ligation with a double-stranded DNA adapter, have been developed [4,6]. The ligation and tailing steps are dependent on the accessibility and lack of base-paired structure at the end of the target cDNA. The recovery of a ligated product is then governed by both the abundance of a particular target cDNA and the net efficiency of adapter ligation and any nucleotide tailing steps.

We sought to ligate an adapter oligonucleotide to single-stranded cDNA-type targets at high efficiencies, independent

of the secondary structure in the DNA. Using a thermostable enzyme with DNA ligase activity [7], we developed an approach for ligating single-stranded DNA targets with complex secondary structures at nearly quantitative efficiencies.

Materials and methods

Oligonucleotides

Oligonucleotides (synthesized by Midland Certified) for cDNA ligation were as follows: DNA adapter, 5'-p-ATCGA CAACA ACTCT CCTCC TCCGT GCG-(C3)-3' (where p is a phosphate and C3 is 1,3-propanediol); DNA target (119 nt), 5'-GGTAA GCTGG CCAGC AACTT ATCTG TGTTT TTCCG ATTGT CTAGT GTCTA TGTTT GATGT TATGC GCCTG CGTCT GTACG TTCTG AACAC CCGGC CGCAA CCCTG GGAGA CGTCC CAGG-3'. PCR primers: forward, 5'-GGTAA GCTGG CCAGC AACTT ATCTG TGTTT TTCCG AT-3'; reverse, 5'-CGCAC GGAGG AGGAG AGTTG TTGTC GAT-3'. Reverse transcriptase primer for structured cDNA, 5'-GAACC GGACC GAAGC CCG-3'. Oligonucleotides for 3' end labeling by single nucleotide extension: complement to DNA target, 5'-AAACC CTGGG ACGTC TCCCA GGGTT GCGG-3'.

* Corresponding author. Fax: +1 919 962 2388.

E-mail address: weeks@unc.edu (K.M. Weeks).

complement to reverse transcriptase primer, 5'-GGGTC GGGCT TCGGT CCGGT TC-3'.

3' ³²P end labeling of DNA oligonucleotides

The DNA target and cDNA reverse transcriptase primer were 3' end labeled using Klenow(-exo) polymerase and annealing to a complementary strand that allows incorporation of a single nucleotide [8]. Reactions (20 μl, 37°C, 40 min) contained 4 μM DNA to be labeled, 4–8 μM complement, 5 U Klenow(-exo) DNA polymerase (New England Biolabs), and 2.5 μM of either α-[³²P]dGTP or α-[³²P]dATP. Reactions were quenched by dilution with formamide and purified by denaturing gel electrophoresis (10–20% [w/v] polyacrylamide, 90 mM Tris–borate, 2 mM ethylenediaminetetraacetic acid [EDTA],¹ 7 M urea). DNAs were recovered from the gel by passive elution at 37°C for 2 h and ethanol precipitation. Target DNA and cDNA reverse transcriptase primer were resuspended in water at final concentrations of 0.1 and 0.3 μM, respectively.

Single-stranded DNA ligation

Ligation reactions (10 μl, 68°C, 1 h) contained 50 mM Mops (pH 7.5), 1 mM dithiothreitol (DTT), 10 mM KCl, 5 mM MgCl₂, 2.5 mM MnCl₂, 20% (w/v) polyethylene glycol 6000 (PEG 6000), 25 μM ATP, ³²P-labeled target DNA, 1 μM adaptor DNA, and 10 U Thermophage DNA ligase (Prokaria). Ligation products were resolved by denaturing electrophoresis in 10% polyacrylamide gels and were quantified using a Phosphorimager (Molecular Dynamics).

Ligation-mediated PCR

Aliquots (8 μl) from each ligation reaction were precipitated with ethanol in the presence of 20 μg glycogen. Pellets were resuspended in 74 μl water, and one-half of this mixture was used in 100-μl PCR reactions (containing 20 mM Tris–HCl [pH 8.4], 50 mM KCl, 0.2 mM each dNTP, 2.5 mM MgCl₂, 0.75 μM forward primer, 0.84 μM reverse primer, 2.5 U *Taq* polymerase). After 32 PCR cycles (94°C, 30 s; 60°C, 30 s; 72°C, 1 min), products were analyzed on 1.4% (w/v) agarose gels.

Construction and analysis of 3' nested set of cDNAs

tRNA transcripts were prepared as described previously [9]. Random 2'-O-adducts were introduced by heating 12 μl of 0.5 μM RNA (in 5 mM Tris [pH 8.0], 0.5 mM EDTA) to 95°C for 1 min, adding 6 μl of 333 mM Hepes (pH 8.0), and treating with 2 μl of 65 mM *N*-methylisatoic anhydride (NMIA) for 1 min at 80°C. Radiolabeled primer was added

(6 μl at 0.3 μM) and annealed at 65°C for 5 min and then at 35°C for 20 min, and primer extension reactions were performed (in 40 μl) exactly as described previously [9]. RNA was subsequently degraded by alkaline hydrolysis (0.2 M NaOH, 95°C, 5 min). The 3' nested set of cDNAs, produced in the primer extension reaction, was initially resolved on a denaturing 20% polyacrylamide sequencing gel. Extension products were identified by comparison with an adjacent sequencing marker, generated by dideoxythymidine nucleotide incorporation during primer extension. cDNAs were then excised from the gel in subsets containing approximately 20 nt each (20–40, 40–60, 60–80, 80–100, 100–120, and 120–132 nt in length). Each size-fractionated subset was individually eluted from the gel (500 mM sodium acetate [pH 6.0], 1 mM EDTA, 37°C, 2 h), concentrated by ethanol precipitation, and resuspended in 10 μl water. Ligation reactions were performed using 1 μl of each cDNA fraction. Ligation products were resolved on 14% denaturing polyacrylamide gels and visualized by phosphorimaging.

Results and discussion

Strategy for optimizing single-stranded DNA ligation

A model system for ligating a DNA adapter to a single-stranded target DNA is shown in Fig. 1. Ligase enzymes in

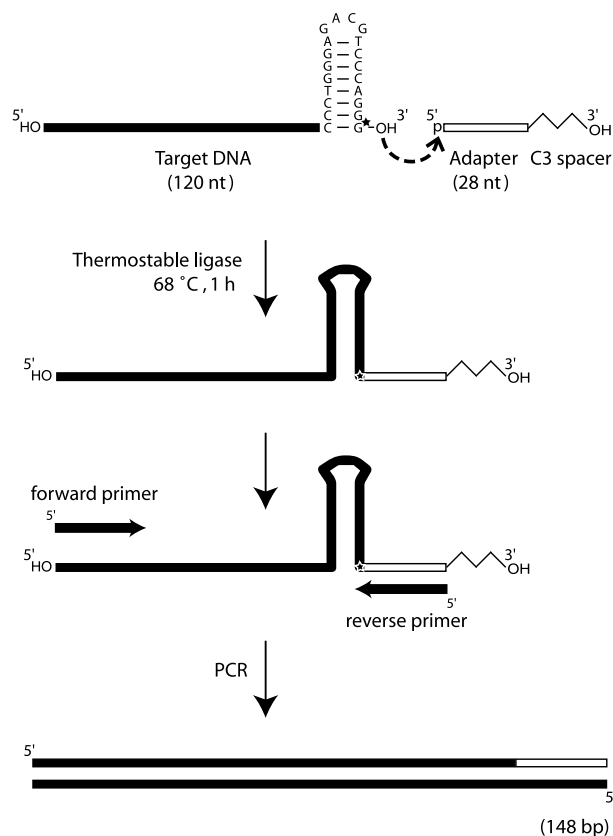


Fig. 1. End-selective single-stranded DNA ligation to create PCR templates. Dashed reaction arrow indicates ligation between 3'-OH and 5'-phosphate (p) groups. Star indicates site of 3' radiolabel used to visualize ligation products (in Fig. 2A).

¹ Abbreviations used: EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol; PEG 6000, polyethylene glycol 6000; NMIA, *N*-methylisatoic anhydride; SHAPE, selective 2'-hydroxyl acylation analyzed by primer extension.

the RNA ligase 1 family catalyze the ATP-dependent formation of a phosphodiester bond between a nucleotide 3'-hydroxyl nucleophile and a free 5'-phosphate group (Fig. 1, dashed reaction arrow) [10]. The adapter oligonucleotide was synthesized with a 5'-phosphate and a nonligatable 3'-C3 (1,3-propanediol) group to enforce unique ligation between the 3'-OH of the target DNA and the 5'-phosphate of the adapter (Fig. 1). The DNA target was radiolabeled at its 3' end and was intentionally designed to form significant secondary structure at its 3' terminus so that, from the onset, the ligation reaction would be optimized against an interfering secondary structure. The adapter oligonucleotide was designed such that it lacks significant secondary structure and simultaneously forms an accessible PCR primer binding site.

Efficiency of single-stranded DNA ligation

Our strategy for efficient single-stranded DNA ligation involves four elements. Reactions were performed (i) at an elevated temperature (68 °C) to reduce secondary structure in the DNA target, (ii) in 20% (w/v) PEG 6000 to increase macromolecular interactions between DNA substrates and the ligase enzyme, (iii) at saturating adapter DNA and ligase concentrations so that ligation efficiency is concentration independent for low amounts of target DNA, and (iv) with the use of an adapter DNA lacking significant self-structure.

Under these optimized conditions, the target DNA was ligated to the adapter oligonucleotide at efficiencies of at

least 97% at DNA concentrations of 0.01–0.1 nM (See ligated product in Fig. 2A). The absence of a low-mobility product in reactions omitting the adapter DNA demonstrates that only linear ligated product, and not DNA circles, form in these reactions (Fig. 2A, see minus adapter lanes). At the higher target DNA concentrations of 1 and 10 nM, ligation efficiencies are 80 and 56%, respectively, and can be increased to at least 90% by extending the reaction time from 1 to 4 h (data not shown). Higher efficiencies at lower target DNA concentrations support the design principle that reactions should be performed at saturating adapter and ligase concentrations.

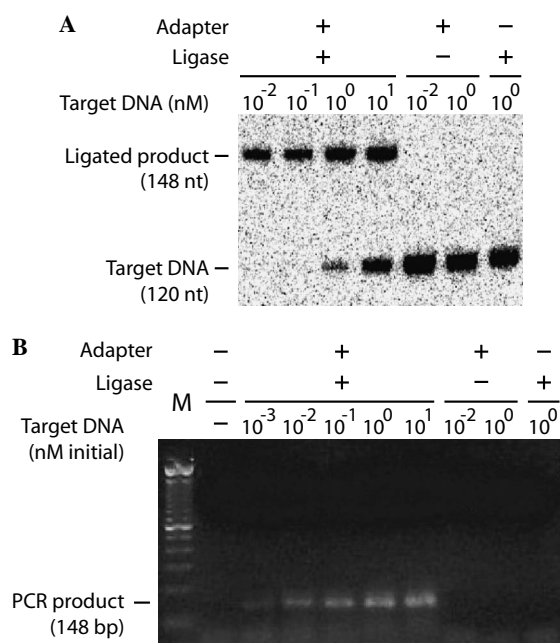


Fig. 2. Efficient single-stranded DNA ligation and selective PCR amplification. (A) Ligation reactions as a function of target DNA concentration (resolved by electrophoresis after adjusting all samples to equal ^{32}P disintegrations). (B) PCR amplification of target DNA adapter ligation products. Marker (M) is a double-stranded DNA ladder in 100-bp increments.

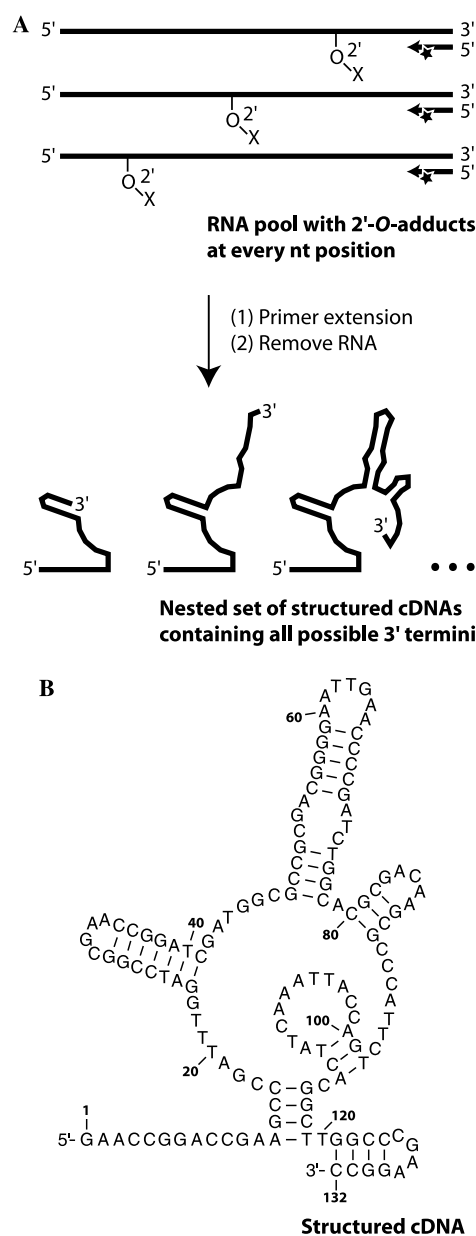


Fig. 3. Construction and secondary structure for cDNA target. (A) Scheme for synthesis of 3' nested set of cDNA fragments. Reverse transcriptase-mediated primer extension was performed using RNA fragments carrying 2'-O-adducts generated by RNA SHAPE chemistry [9]. Site of ^{32}P radiolabel in reverse transcription primer is indicated with a star. (B) Calculated secondary structure for full-length cDNA target [12].

PCR amplification of ligated single-stranded DNAs

The products of the single-stranded DNA ligation can be used directly for PCR (Fig. 1). Ligation reactions (Fig. 2A) were recovered by ethanol precipitation and used without further purification in PCR (Fig. 2B). After 32 cycles, a single product is clearly observable by PCR for the ligated product formed from 10^{-2} nM starting target DNA (corresponding to 0.04 fmol DNA). PCR product synthesis, as expected, requires both the DNA adapter and ligase enzyme (Fig. 2B, control lanes). Thus, using an optimized approach for ligation of single-stranded DNAs, we form target DNA adapter products that are efficiently and selectively amplified by PCR.

Single-stranded DNA ligation independent of target DNA structure

We then evaluated the sensitivity of the single-stranded DNA ligation reaction to the secondary structure present at the 3' end of a target DNA. We generated a set of DNAs with a common 5' end as a complete set of nested 3' extension products by performing a primer extension reaction using a previously described RNA template [9] (Fig. 3A). The RNA template was subjected to RNA selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) chemistry [9,11] under denaturing conditions to yield 2'-O-

adducts at every internal RNA residue. These 2'-O-adducts prevent reverse transcriptase-mediated primer extension beyond the site of the adduct to yield a complete set of cDNA fragments containing every possible DNA from the 19-nt long primer to the 132-nt full-length product (Fig. 3A). Although there is some variation in individual bands, all expected extension products are observed as bands of significant intensity in a sequencing gel (Fig. 4, complete lane).

There are 113 different DNA products in the complete nested set, and the 3' ends of these products lie in both highly structured stem loops and unpaired regions (Fig. 3B). For analysis, the complete nested set was divided into subsets containing approximately 20 DNA fragments each by excising and purifying cDNA fragments from a preparative gel (Fig. 4). Reactions were performed in the presence (+) and absence (–) of the thermostable ligase and were resolved on sequencing gels. For each size-fractionated cDNA subset, ligation to the adapter oligonucleotide is detected as the disappearance of a band in the (–) ligase lane and a shift to a product 28 nt longer in length in the (+) ligase lane. For clarity, products longer than 60 nt were subjected to extended electrophoresis to achieve improved resolution (Fig. 4, right panel).

In nearly every case, target DNAs are ligated at greater than 95% efficiency. For example, in the reactions of DNAs 40–60, 60–80, 80–100, and 100–120 nt in length, all bands corresponding to the starting DNAs are quantitatively

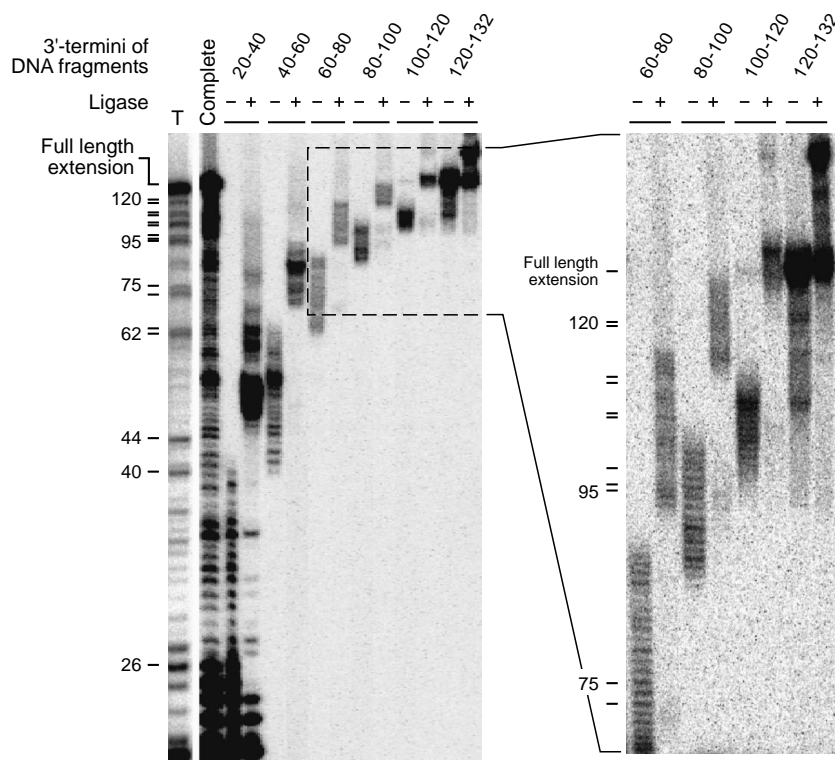


Fig. 4. Structure-independent ligation of nested structured cDNAs. Complete lane shows initial 3' nested set of cDNA fragments; size-fractionated sets shown in remaining lanes were created from this parent ensemble. All cDNAs have identical 5' ends but differ in length, as shown. cDNA fractions were incubated with the 28-nt adapter oligonucleotide in the presence (+) or absence (–) of thermostable ligase. Sequencing marker (T) was generated by incorporation of dideoxythymidine during primer extension from the RNA template; positions of thymidine residues are indicated at the left of each panel. Right-hand panel shows extended electrophoresis of selected reactions from main panel.

converted to product DNAs (compare – and + ligase lanes in Fig. 4). The main exception to quantitative ligation is the experiment involving DNAs 120–132 nt in length, where the full-length extension product was ligated to approximately 50%; most shorter cDNAs in this group were ligated at more than 80% efficiency.

Conclusion

Using a combination of a thermostable ligase, elevated temperature, PEG-induced molecular exclusion, and the forcing of product formation with saturating ligase and adapter DNA concentrations, we achieved efficient ligation of single-stranded DNAs independent of secondary structure in the DNA target. This approach is directly compatible with, and should yield immediate improvements to, many current strategies for analyzing RNA structure via cDNA intermediates. Efficient structure-independent cDNA ligation also represents an important technology for our ongoing program directed toward quantitative analysis of RNA structure in vivo.

Acknowledgments

This work was supported by grants from the University of North Carolina at Chapel Hill Office of Undergraduate Research and the Frances C. and William P. Smallwood Foundation (to Tian Li) and from the National Science Foundation (MCB-0416941 to Kevin Weeks). We are indebted to Kevin Wilkinson for assistance with RNA SHAPE chemistry.

References

- [1] X. Liu, M.A. Gorovsky, Mapping the 5' and 3' ends of *Tetrahymena thermophila* mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE), *Nucleic Acids Res.* 21 (1993) 4954–4960.
- [2] J.B. Edwards, J. Delort, J. Mallet, Oligodeoxyribonucleotide ligation to single-stranded cDNAs: a new tool for cloning 5' ends of mRNAs and for constructing cDNA libraries by in vitro amplification, *Nucleic Acids Res.* 19 (1991) 5227–5232.
- [3] E. Bertrand, M. Fromont-Racine, R. Pictet, T. Grange, Visualization of the interaction of a regulatory protein with RNA in vivo, *Proc. Natl. Acad. Sci. USA* 90 (1993) 3496–3500.
- [4] H.H. Chen, D. Castanotto, J.M. LeBon, J.J. Rossi, A.D. Riggs, In vivo, high-resolution analysis of yeast and mammalian RNA–protein interactions, RNA structure, RNA splicing, and ribozyme cleavage by use of terminal transferase-dependent PCR, *Nucleic Acids Res.* 28 (2000) 1656–1664.
- [5] D.C. Tessier, R. Brousseau, T. Vernet, Ligation of single-stranded oligodeoxynucleotides by T4 RNA ligase, *Anal. Biochem.* 158 (1986) 171–178.
- [6] W.M. Schmidt, M.W. Mueller, Controlled ribonucleotide tailing of cDNA ends (CRTC) by terminal deoxynucleotidyl transferase: a new approach in PCR-mediated analysis of mRNA sequences, *Nucleic Acids Res.* 24 (1996) 1789–1791.
- [7] T. Blondal, A. Thorisdottir, U. Unnsteinsdottir, S. Hjorleifsdottir, A. Aevarsson, S. Ernstsson, O.H. Fridjonsson, S. Skirnisdottir, J.O. Wheat, A.G. Hermannsdottir, S.T. Sigurdsson, G.O. Hreggvidsson, A.V. Smith, J.K. Kristjansson, Isolation and characterization of a thermostable RNA ligase I from a *Thermus scotoductus* bacteriophage TS2126 with good single-stranded DNA ligation properties, *Nucleic Acids Res.* 33 (2005) 135–142.
- [8] Z. Huang, J.W. Szostak, A simple method for 3'-labelling of RNA, *Nucleic Acids Res.* 24 (1996) 4360–4361.
- [9] E.J. Merino, K.A. Wilkinson, J.L. Coughlan, K.M. Weeks, RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE), *J. Am. Chem. Soc.* 127 (2005) 4223–4231.
- [10] O.C. Uhlenbeck, R.J. Gumport, RNA ligase, *Enzymes* 15 (1982) 31–58.
- [11] K.A. Wilkinson, E.J. Merino, K.M. Weeks, RNA SHAPE chemistry reveals nonhierarchical interactions dominate equilibrium structural transitions in tRNA^{ASP} transcripts, *J. Am. Chem. Soc.* 127 (2005) 4659–4667.
- [12] D.H. Mathews, M.D. Disney, J.L. Childs, S.J. Schroeder, M. Zuker, D.H. Turner, Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure, *Proc. Natl. Acad. Sci. USA* 101 (2004) 7287–7292.