ORGANIC LETTERS

2003
Vol. 5, No. 3
273−276

Synthesis of Lariat-DNA via the Chemical Ligation of a Dumbbell Complex

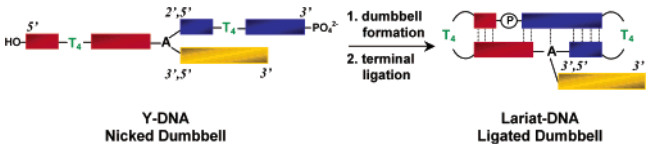
Sandra Carriero and Masad J. Damha*

Department of Chemistry, McGill University, Montreal, Quebec, Canada H3A 2K6

masad.damha@mcgill.ca

Received November 4, 2002

ABSTRACT



An efficient synthesis of a medium-sized DNA lariat through the chemical ligation of a Y-shaped dumbbell precursor is described. The methodology requires only commercially available phosphoramidites and reagents and affords regioisomerically pure lariat molecules. Characterization of the lariat by T_m analysis reveals that the molecule displays markedly enhanced thermal stability and unimolecular association–dissociation kinetics, consistent with DNA dumbbell behavior.

"Lariat" RNAs have been of specific interest to both the chemical and biological communities since their discovery nearly two decades ago as splicing intermediates (branched RNA introns) in the biosynthesis of mRNA.[1] The inherent branched architecture of these molecules is ascribed to a highly conserved branch point adenosine moiety bearing vicinal 2′,5′- and 3′,5′-internucleotidic phosphodiester linkages.[2] Synthetic branched analogues (Y-shaped molecules) of these unique biomolecules have been particularly useful for probing the structural requirements of branch point recognition during splicing[3] and of lariat debranching enzymes.[4] In addition, they have been used as haptens for isolating antibodies specific to branched oligonucleotides,[5] as potential agents for controlling gene expression,[6] and as nucleic acid biosensors.[7] Although many groups, including our own, have been successful in chemically synthesizing branched oligonucleotides (Y-shaped),[8] synthesis of nucleic acid lariats has proven to be a major challenge. RNA "mini"-lariat syntheses have been reported;[9] nevertheless, the methodologies require the use of nonstandard nucleoside building blocks, are extensively time-consuming (many chromatographic purification steps), and are restricted to the synthesis of small lariats (trimeric and pentameric loops), which are nonrepresentative of the naturally occurring biomolecules.

We report here an effective and versatile strategy for synthesizing lariat DNA larger than those that have previously been synthesized by exploiting the intramolecular

(1) (a) Wallace, J. C.; Edmonds, M. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 950−954. (b) Padgett, R. A.; Konarska, M. M.; Grabowski, P. J.; Hardy, S. F.; Sharp, P. A. *Science* **1984**, *225*, 898−903. (c) Ruskin, B.; Krainer, A. R.; Maniatis, T.; Green, M. R. *Cell* **1984**, *38*, 317−331.

(2) Domdey, H.; Apostol, B.; Lin, R.-J.; Newman, A.; Brody, E.; Abelson, J. *Cell* **1984**, *39*, 611−621.

(3) Carriero, S.; Braich, R. S.; Hudson, R. H. E.; Anglin, D.; Friesen, J. D.; Damha, M. J. *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 873−877.

(4) Nam, K.; Hudson, R. H. E.; Chapman, K. B.; Ganeshan, K.; Damha, M. J.; Boeke, J. D. *J. Biol. Chem.* **1994**, *269*, 20613−20621.

(5) Reilly, J. D.; Freeman, S. K.; Melhem, R. F.; Kierzek, R.; Caruthers, M. H.; Edmonds, M.; Munns, T. W. *Anal. Biochem.* **1990**, *185*, 125−130.

(6) (a) Hudson, R. H. E.; Ganeshan, K.; Damha, M. J. *Branched Oligonucleotides: Synthesis and Biological Applications*; Cook, P. D., Sanghvi, Y. S., Eds.; American Chemical Society: Washington, DC, 1994; Vol. 580, pp 133−152. (b) von Buren, M.; Petersen, G. V.; Rasmussen, K.; Brandenburg, G.; Wengel, J. *Tetrahedron* **1995**, *51*, 8491−8506.

(7) Uddin, A. H.; Piunno, P. A. E.; Hudson, R. H. E.; Damha, M. J.; Krull, U. J. *Nucleic Acids Res.* **1997**, *25*, 4139−4146.

(8) (a) Sekine, M.; Hata, T. *J. Am. Chem. Soc.* **1985**, *107*, 5813−5815. (b) Damha, M. J.; Pon, R. T.; Ogilvie, K. K. *Tetrahedron Lett.* **1985**, *26*, 4839−4842. (c) Damha, M. J.; Zabarylo, S. *Tetrahedron Lett.* **1989**, *30*, 6295−6298. (d) Grotli, M.; Eritja, R.; Sproat, B. *Tetrahedron* **1997**, *53*, 11317−11346.

(9) (a) Sund, C.; Agback, P.; Chattopadhyaya, J. *Tetrahedron* **1991**, *47*, 9659−9674. (b) Sund, C.; Agback, P.; Chattopadhyaya, J. *Tetrahedron* **1993**, *49*, 649−668. (c) Reese, C. B.; Song, Q. *Nucleic Acids Res.* **1999**, *27*, 2672−2681.

10.1021/ol027227e CCC: $25.00
Published on Web 01/09/2003

hybridization characteristics of nucleic acid "dumbbells". Dumbbells (i.e., double-hairpin structures) consist of an oligonucleotide duplex that is covalently closed on both ends by nonhybridizing loop structures (nucleotidic or synthetic linkers) and typically undergo unimolecular order−disorder transitions.[10] The synthesis of linear DNA and chimeric DNA/RNA "dumbbells" is well established and is accomplished by the intramolecular condensation of a properly aligned proximal phosphate and hydroxyl group by either chemical[11] or more costly enzymatic means (T4 DNA or RNA ligase),[12] albeit this is the first account of an intramolecular ligation of a branched DNA molecule. Before tackling DNA lariat synthesis, the optimal dumbbell precursor had to be designed. This was accomplished by synthesizing a series of linear DNA dumbbells via solid-phase phosphoramidite chemistry in order to evaluate the following: (1) the effect of incorporating a 2′,5′-riboadenosine (2′-rA) insert and dT_4 hairpin loop on the thermal stability of the complex; (2) the ligation efficiency of a 3′-phosphate/5′-OH versus a 5′-phosphate/3′-OH; (3) the optimal nucleotide composition at the ligation junction, and (4) the effectiveness of ligation directly opposite or distal to the 2′,5′-rA insert (Table 1, sequences **1**−**9**). The initial dumbbell sequence (**3**−**6**) was chosen so that comparisons could be made with a similar complex previously studied by Ashley and Kushlan, which merely lacked the 2′-rA insert.[13] All molecules were designed to favor exclusive formation of intramolecular duplexes.

Table 1. List of Sequences Used in Dumbbell Study

DNA	oligonucleotide sequence (5'→3')[a]
Bimolecular Complexes	
1	^{HO}CGC-rA$_{3'5'}$-CGC$_{OH}$
2	^{HO}CGC-rA$^{2'5'}$-CGC$_{OH}$
Unimolecular (Dumbbell) Complexes	
3	^{HO}GCG-T$_4$-CGC-rA$_{3'5'}$-CGC-T$_4$-GCGT$_{OH}$
4	^{HO}GCG-T$_4$-CGC-rA$^{2'5'}$-CGC-T$_4$-GCGT$_{OH}$
5	^{P}GCG-T$_4$-CGC-rA$^{2'5'}$-CGC-T$_4$-GCGT$_{OH}$
6	^{HO}GCG-T$_4$-CGC-rA$^{2'5'}$-CGC-T$_4$-GCGT$_{P}$
7	^{HO}TGCG-T$_4$-CGCA-rA$^{2'5'}$-CGC-T$_4$-GCGT$_{P}$
8	^{HO}TGTTGCG-T$_4$-CGCAACA-rA$^{2'5'}$-CGC-T$_4$-GCGT$_{P}$
9	^{HO}TGCG-T$_4$-CGCAACA-rA$^{2'5'}$-CGC-T$_4$-GCGTTGT$_{P}$
10	^{HO}TGCG-T$_4$-CGCAACA-rA$^{2'5'}$-CGC-T$_4$-GCGTTGT$_{P}$ (3′5′ → T$_{16}$)

[a] Sequences represent the open (nicked) forms of the DNA dumbbells synthesized. P = phosphate monoester group.

Effect of the T_4 Loop and a 2′,5′-Riboadenosine Substitution on DNA Dumbbell Stability. It has previously been shown by our group[14] and others[15] that the incorporation of 2′,5′-RNA constituents into and otherwise unmodified oligonucleotide sequence results in a decrease in the melting temperature (T_m) of the duplex. Using thermal melting analysis, we evaluated the effect of a single 2′-rA insertion on the stability of DNA dumbbell **4** and compared it to its homologous counterpart **3** containing a 3′-rA (Table 2). The inclusion of one 2′-rA unit in dumbbell **4** resulted in destabilization of the complex by 3 °C; however, in the complex lacking the T_4 closing loop, this substitution resulted in a decrease in T_m of 5 °C (compare **1** and **2**).[16] This implies that the homonucleotide loop cooperatively stabilizes the dumbbell structures, thereby sustaining the 2′-rA substitution to a greater extent than the linear duplexes.[17] The resultant T_m of **4** (50 °C) establishes that the 2′-rA insertion is well tolerated in the dumbbell state, allowing us to proceed with investigating its ligation behavior.

Table 2. Summary of T_m Values for Complexes[a]

DNA chimera	T_m of DNA complex (°C) A nicked	T_m of DNA complex (°C) B ligated	nucleotides at nicked junction	% yield ligated product[b]
1[c]	35			
2[c]	30			
3	53			
4	50			
5	50	>85	GpT	4[d]
6		>85	GpT	10[d]
7	54	87	TpT	66[e]
8	62	88	TpT	79[e]
9	54	88	TpT	88[e]
10	52	86	TpT	37[f]

[a] T_m buffer: 10 mM Tris-HCl, 10 mM NaCl pH 7.5. Temperatures are averages of two successive runs and are within ±1°C. [b] % yield is the ratio of ligated dumbbell to unreacted precursor. [c] T_m values of **1** and **2** were obtained by hybridization to their corresponding DNA complement: 5′-GCGTGCG-3′. [d] Compounds **5** and **6**: Ligations were performed with *N*-cyanoimidazole and the amount of ligated product determined by densitometry (UN-SCAN-IT Software: Silk Scientific, Utah, 1996). [e] Compounds **7**−**9**: Ligations were performed with CNBr, and the extent of ligation was determined by HPLC integration of the product peak. [f] Compound **10**: Ligation was performed with CNBr and the amount of ligated product determined by densitometry.

Optimization of Dumbbell Ligation Efficiency. Terminal 5′- or 3′-phosphate groups were introduced into nicked precursors using the commercially available phosphitylating reagent developed by Horn and Urdea.[18] Ligation of all DNA dumbbells were conducted at a concentration of 10^{-4} M phosphorylated nicked precursor (**5**−**9**) using either *N*-

(10) Wemmer, D. E.; Benight, A. S. *Nucleic Acids Res.* **1985**, *13*, 8611−8621.

(11) (a)Ashley, G. W.; Kushlan, D. M. *Biochemistry* **1991**, *30*, 2927−2933. (b) Kuznetsova, S. A.; Merenkova, I. N.; Kanevsky, I. E.; Shabarova, Z. A.; Blumenfeld, M. *Antisense Nucleic Acid Drug Dev.* **1999**, *9*, 95−100.

(12) (a) Erie, D.; Sinha, N.; Olson, W.; Jones, R.; Breslauer, K. *Biochemistry* **1987**, *26*, 7150−7159. (b) Park, W.-S.; Miyano-Kurosaki, N.; Abe, T.; Takai, K.; Yamamoto, N.; Takaku, H. *Biochem. Biophys. Res. Commun.* **2000**, *270*, 953−960.

(13) See ref 11a. DNA dumbbell studies by Ashley and Kushlan: 5′-GCG-T$_4$-CGCCGC-T$_4$-GCG-3′: T_m(nicked) = 58 °C; T_m(ligated) = 86 °C.

(14) (a)Giannaris, P. A.; Damha, M. J. *Nucleic Acids Res.* **1993**, *21*, 4742−4749. (b) Wasner, M.; Arion, D.; Borkow, G.; Noronha, A.; Uddin, A. H.; Parniak, M. A.; Damha, M. J. *Biochemistry* **1998**, *37*, 7478−7486.

(15) Kandimalla, E. R.; Manning, A.; Zhao, Q.; Shaw, D. R.; Byrn, R. A.; Sasisekharan, V.; Agrawal, S. *Nucleic Acids Res.* **1997**, *25*, 370−378.

(16) See ref 14b. One insert of 2′-rA into an rA$_{10}$ homopolymer resulted in a decrease of −4.2 °C/insert when bound to poly-dT.

(17) Bloomers, M. J. J.; Walters, J. A. L. I.; Haasnoot, C. A. G.; Aelen, J. M. A.; van der Marel, G. A.; van Boom, J. H.; Hilbers, C. W. *Biochemistry* **1989**, *28*, 7491−7498.

(18) Horn, T.; Urdea, M. S. *Tetrahedron Lett.* **1986**, *27*, 4705−4708.

cyanoimidazole (CNIm)[19] or cyanogen bromide (CNBr)[20] as phosphodiester condensing reagents (see Supporting Information). T_m analysis of the nicked and ligated dumbbells demonstrated cooperative monophasic transitions. Ligation of all compounds resulted in the formation of a single new product band, which migrated faster than the corresponding nicked precursor on a denaturing polyacrylamide gel (Figure 1). Conceivably, this is the result of the more globular and compact nature of the circularized products. A comparison of the T_m profiles of the nicked and ligated dumbbells reveals that the ligated molecules demonstrate a significant enhancement (≥30 °C) in melting temperature (Table 2). This increased thermal stability is consistent with previous studies on dumbbell structures and allows us to utilize this as a means of characterization.[21] All nicked precursors and ligated dumbbells were further characterized by negative-ion mode MALDI-MS.[22]

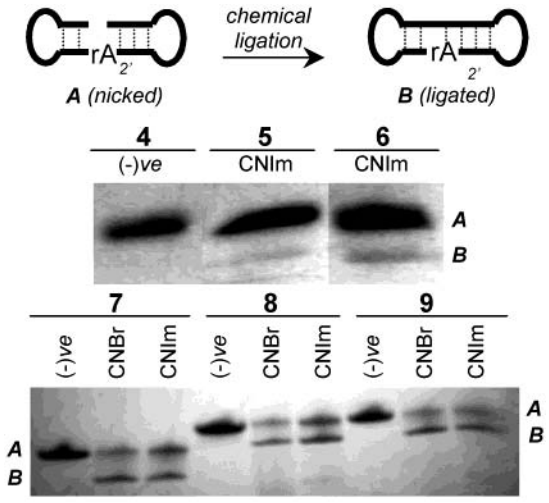


Figure 1. 20% Denaturing (8.3 M urea) PAGE of linear DNA dumbbell ligation reactions conducted using either CNIm or CNBr reagent or both. A: nicked dumbbell precursor. B: ligated dumbbell.

Localization of the phosphate group participating in ligation (5′-end or 3′-end) has been described as a key factor influencing the efficacy of chemical ligation. Our findings corroborate previous reports that ligation yield improves when the precursor contains a 5′-hydroxyl and a 3′-phosphate;[11a,23] however, the yields obtained were significantly lower than desired (Table 2 and Figure 1: **5** and **6**). Moreover, it was revealed to us that the chemical ligation efficiency also strongly depends on the nature of the nucleotide residues to be ligated, with TpT bearing the most productive contact at the ligation site.[11b,24] When the sequence was modified such that the nucleotide junction contained a TpT (**7**) rather than a GpT contact (**6**), the ligation yield increased by ca. 56% (Table 2). It is worth mentioning that all the ligations described so far were conducted at a nicked junction directly opposite the 2′-rA insertion. We imagined that this slightly destabilizing insertion may be altering the local base-stacking interactions surrounding the junction, and as such, the ligation efficiency should improve if the TpT junction were placed at a site remote from the 2′-rA substitution. Sequences **8** and **9** were designed to contain identical base compositions so that we could compare the effect of ligation directly across from (**8**) or distal to (**9**) the 2′-rA insert. Furthermore, to accommodate this change in sequence design, additional base pairs were introduced into the stems of the dumbbells. The new sequences now had [7 + 4] nucleotides complementary in the template (compared to [4 + 4] nucleotides for **7**), thereby increasing the thermal stability of the double-hairpin complexes. This appears to be the case for **8**, whereby the T_m increases by 8 °C compared to **7**, yet the T_m of **9** remains unchanged (54 °C). These results suggest that the 2′-rA has a more detrimental destabilizing effect at the center of a duplex region (**9**) rather than at the end (**8**). While this may be the case, it seems that this scenario can be used to our advantage. Since the local geometry of the duplex appears to be distorted in the region encompassing the 2′-rA unit, then better base-stacking should be achieved further away. Indeed this is what we observed, and the ligation yield improved by as much as 10% for compound **9** as compared to **8**, signifying an improvement in the stacking interaction and base-pair alignment of the TpT junction (Figure 1, Table 2). Furthermore, the nature of the condensing reagent, either *N*-cyanoimidazole or cyanogen bromide, did not appear to be an important factor in optimizing the ligation efficiency, as the extent of reaction is similar for both reagents (Figure 1). Nevertheless, CNBr-induced ligations are clearly advantageous because the reaction time is substantially decreased (5 min vs 4 h). Since cyanogen bromide is known to be a highly toxic chemical, we recommend that the choice of condensing reagent be made at the discretion of the experimenter.

Synthesis of Lariat DNA. The approach we adopted for lariat DNA synthesis uses commercially available phosphoramidite building blocks in conjunction with our previously published protocol for the regiospecific solid-phase synthesis of branched DNA molecules,[25] phosphitylation of the 3′-terminus to yield a 3′-phosphate, and chemical condensation of the nicked phosphate/hydroxyl junction using CNBr (see Supporting Information for scheme). The lariat precursor **10** was designed such that the 5′- and 2′-extensions of the branched (Y-shaped) precursor are capable of intramolecularly folding into a "dumbbell" complex, thereby aligning the reactive 5′-OH and 3′-phosphate for chemical

(19) Ferris, J. P.; Huang, C.-H.; Hagan, W. J., Jr. *Nucleosides Nucleotides* **1989**, *8*, 407−414.
(20) Shabarova, Z. A.; Federova, O. A.; Dolinnaya, N. G.; Gottikh, M. B. *Orig. Life Evol. Biochem.* **1997**, *27*, 555−566.
(21) Erie, D. A.; Jones, R. A.; Olson, W. K.; Sinha, N. K.; Breslauer, K. J. *Biochemistry* **1989**, *28*, 268−273.
(22) Distler, A. M.; Allison, J. *Anal. Chem.* **2001**, *73*, 5000−5003.
(23) Shabarova, Z. A. *Biochimie* **1988**, *70*, 1323−1334.
(24) Merenkova, I. P.; Dolinnaya, N. G.; Oretskaia, T. S.; Sokolova, N. I.; Shabarova, Z. A. *Bioorg. Khim.* **1992**, *18*, 85−91.
(25) (a)Braich, R. S.; Damha, M. J. *Bioconj. Chem.* **1997**, *8*, 370−377. (b) Carriero, S.; Damha, M. J. *Solid-Phase Synthesis of Branched Oligonucleotides*; Beaucage, S., Ed.; Wiley & Sons: 2002; Vol. 4.14, pp 1−32.

ligation. The 5′- and 2′-extensions (dumbbell portion) contained a base sequence identical to that of **9**, as this was determined to be our optimal substrate for ligation. The 3′ extension of the molecule was constructed so that it did not share sequence complementarity with any part of the 5′- or 2′-extensions and, as such, is excluded from the dumbbell complex. This allows us to circumvent the formation of the undesired 3′,5′-linked lariat regioisomer. We anticipated that such a dumbbell complex would have significant potential to form because previous experiments on branched molecules have indicated a preferred base-stacking interaction between the branch point adenine and the 2′,5′-linked substituent, which stabilizes the molecule.[26] T_m measurement of the hybridized nicked lariat precursor **10** exhibited a monophasic transition (Figure 2B) and displayed thermal stability similar

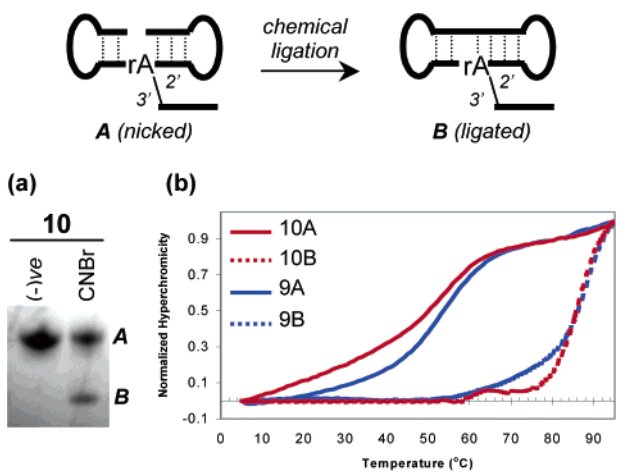


Figure 2. Lariat DNA dumbbell ligation. (a) 12% denaturing PAGE (8.3 M urea) of lariat ligation using CNBr. (b) Characterization of dumbbell lariat structure by T_m analysis and comparison to the corresponding dumbbell structure lacking the 3′-extension (**9**).

to that of its linear nicked complex **9** (ΔT_m = 2 °C). This indicates that the 3′-extension of the molecule (3′-dT_{16}) does not alter the helix geometry of the dumbbell complex by protruding into the solvent, in accordance with a preferential 2′,5′-stacking interaction.[26] Ligation of the lariat dumbbell **10** proceeded using the same reaction conditions described above for the cyclization of linear dumbbells and resulted in the formation of a single, faster moving product as evidenced by denaturing PAGE (Figure 2A). Confirmation of the desired lariat was accomplished by negative MALDI-MS (see Supporting Information) and enhancement in thermal stability of the ligated complex compared with the nicked (ΔT_m = 34 °C; Figure 2B). We believe that the yield of lariat DNA **10** was lower than anticipated (relative to **9**) due to the increased sterics of extending the 2′-chain in the vicinity of the solid support during Y-DNA synthesis, resulting in incomplete phosphitylation of the 3′-terminus of the Y-DNA precursor (Table 2). This molecule was purified by denaturing PAGE and, as such, it is possible that the 3′-phosphorylated product comigrated with some of the nonphosphorylated species, which cannot be cyclized (n-1 band). An alternative approach would be to use a functionalized solid matrix that releases a 3′-phosphorylated oligonucleotide at the end of a synthesis yet can withstand the intermediary deprotection conditions required for the regiospecific Y-DNA synthesis.[27]

Our results demonstrate the first example of the synthesis of a medium-sized DNA lariat oligonucleotide, comprised of a 30-nucleotide loop and a 16-nucleotide stem. Furthermore, we have been able to show that a branched DNA molecule is capable of folding into a lariat−dumbbell complex without distorting the helix geometry of the intramolecular duplex. This hybridization/ligation methodology can potentially allow for the synthesis of the more biologically relevant RNA lariats from branched RNA precursors. Although the construction of such lariats is restricted to those precursors that contain internal templates for hybridization, we are currently investigating the possibility of synthesizing lariats using an intermolecular template (bimolecular reaction), thereby allowing us to synthesize lariats of any sequence composition. Such molecules would be of interest for potentially targeting gene expression, as nucleic acid “aptamers” and biomolecular nanodevices.

Acknowledgment. We thank NSERC for financial support. S.C. thanks FCAR for a postgraduate fellowship. Thanks are due to M. Mangos and L. Fader for proofreading the manuscript.

Supporting Information Available: Experimental details, MALDI-TOF-MS data, and anion-exchange HPLC profiles. This material is available free of charge via the Internet at http://pubs.acs.org.

OL027227E

(26) (a)Damha, M. J.; Ogilvie, K. K. *Biochemistry* **1988**, *27*, 6403−6416. (b) Sund, C.; Agback, P.; Koole, L. H.; Sandstrom, A.; Chattopadhyaya, J. *Tetrahedron* **1992**, *48*, 695−718.

(27) Markiewicz, W. T.; Wyrzykiewicz, T. K. *Nucleic Acids Res.* **1989**, *17*, 7149−7158.