Journal Name

Dynamic Article Links ►

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Oligonucleotide cyclization: The thiol-maleimide reaction revisited

Albert Sánchez,[a] Enrique Pedroso[a] and Anna Grandas*[a]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

A novel method to synthesize cyclic oligonucleotides (5- to 26-mer) using the thiol-maleimide reaction is described. The target molecules were obtained after subsequent removal of thiol and maleimide protecting groups from 5'-maleimido-3'-thiol-derivatized linear precursors. Retro-Diels–Alder conditions deprotecting the maleimide simultaneously promoted cyclization cleanly and in high yield.

The advantageous properties of cyclic oligonucleotides with respect to linear counterparts (stability to exonucleases, high affinity and selectivity in binding target sequences) have turned these analogs into tools useful for diagnostics and inhibition of gene expression (1). Circularity imposes conformational restrictions and reduces fraying of internal associations, which offers opportunities for assessing enzyme processivity in highly demanding conditions (2) as well as for structural studies (3). Furthermore, the discovery that small cyclic oligonucleotides are involved in bacterial metabolism has triggered interest in these molecules (4).

Even though various procedures have been reported, the synthesis of cyclic oligonucleotides is still a challenging goal. Cyclization reactions have sometimes been carried out on a solid support after chain assembly (5) or during the deprotection step (6), but most often in solution, either chemically or enzymatically. An external template placing the two ends close to each other has generally been used to obtain circles with more than 30 nucleotides (7). Preorganization of linear precursors as hairpins or in multi-stranded structures involving intramolecular interactions (i-motif, G-quadruplexes) has also been exploited to favor proximity between the two ends (8). Regarding the covalent union and depending on the methodology used, circularization has furnished either natural phosphodiesters or unnatural linkages (6, 8f, 9). In the past few years, where the number and variety of applications of so-called click reactions (10) has exponentially increased, different groups have used the Huisgen cycloaddition to synthesize cyclic oligonucleotides with a triazole moiety (8f,11).

In this manuscript we wish to describe, for the first time, that cyclic oligonucleotides can also be obtained using the Michael-type addition of thiols to maleimides, taking advantage of the methodology that allows maleimido-oligonucleotides to be on-resin assembled (12). The key point of this methodology is the use of a maleimide protecting group that remains unaltered during the ammonia-promoted deprotection of oligonucleotide functional groups, and which can be subsequently eliminated through a retro–Diels-Alder reaction.

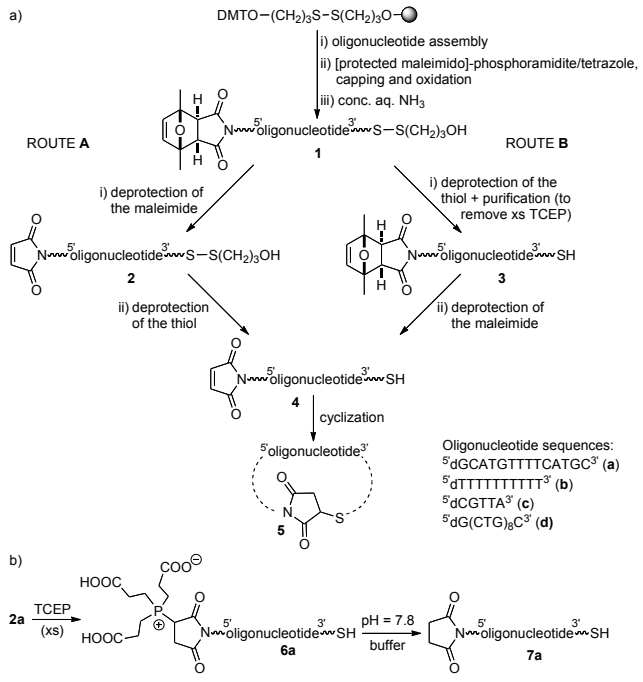
Assessment of whether the thiol-maleimide reaction was suitable for oligonucleotide cyclization required synthesis of a linear oligonucleotide precursor containing masked maleimide and thiol functionalities at either end (**1**, Scheme 1a). Then, two different deprotection sequences were examined (Routes A and B, Scheme 1a) to obtain the target cyclic molecule **5**. First experiments were carried out with oligonucleotide sequence $^{5'}$dGCATGTTTTCATGC (**a**), which in principle could form either a hairpin or a duplex with an internal loop (Figure S4).

Linear precursor **1a** (Scheme 1a) was assembled on a support enabling thiol-derivatization of the 3' end (Scheme S1). Chain elongation was followed by incorporation of the protected maleimide at the 5' end using a phosphoramidite derivative now also commercially available. Then, treatment with concd. aqueous ammonia at room temperature deprotected the oligonucleotide functionalities, leaving the maleimide and thiol protecting groups unaltered.

As outlined in Route A, the maleimide moiety of oligonucleotide **1a** was first deprotected (heat-promoted retro-Diels–Alder reaction) to afford partially protected oligonucleotide **2a**. Reduction of the disulfide cannot be carried out with a thiol, because it would react with the free maleimide. Water-soluble triscarboxyethylphosphine (TCEP) was chosen as reducing agent. Reaction of **2a** with a large excess of TCEP (1000 equiv) did not provide oligonucleotide **4a** (Scheme 1a). The main product in the crude was found to have a mass corresponding to [**4a** + $P(CH_2CH_2COOH)_3$].

Phosphines, and thus TCEP, are known to exhibit a relatively low reactivity towards maleimides, much lower than thiols (13), but addition can indeed take place. Taking into account that the presence of a proton donor has been described to favor formation of fairly stable zwitterionic species (14), a plausible structure for the first product formed is **6a** (Scheme 1b). Zwitterion **6a** could also form a neutral species with a pentacoordinated phosphorus. In slightly basic aqueous solution (conditions suitable for thiol-maleimide reaction, pH=7.8) this product (**6a**) was transformed into succinimide **7a** (as assessed by mass spectrometric analysis). Reduction to succinimide could plausibly take place as described for the reaction of maleimides with triphenylphosphine in MeOH (15).

Use of a lower excess of TCEP (10 equiv) did not provide **4a**, but a side product whose mass was consistent with a non reduced disulfide accompanied by TCEP addition to the maleimide. Consequently, route A was abandoned to explore route B.



Scheme 1. a) Alternatives for the preparation of cyclic oligonucleotides from linear precursors **1**. b) Undesired reaction of maleimides with TCEP, and reduction to succinimide.

Attempts to obtain **4a** from **1a** in a one-pot experiment (that is, thiol deprotection followed by maleimide deprotection without any purification) were unsuccessful. Treatment of **1a** with TCEP afforded **3a**, but the crude obtained after subsequent maleimide deprotection contained a mixture of **6a** and **7a** (see structures in Scheme 1b). Conversely, when disulfide reduction of **1a** was followed by HPLC purification to remove excess TCEP, maleimide deprotection of pure **3a** afforded a crude in which the main compound (55-60 % by HPLC, data not shown) exhibited an *m/z* ratio corresponding to either **4a** or **5a**.

This compound remained unaltered after incubation in conditions optimal for cyclization (aqueous buffer, slightly basic pH). Since in aqueous media maleimido-oligonucleotides are easily transformed into products incorporating a water molecule, this result suggested that the final compound was likely to be the cyclic molecule, **5a**. It is interesting to notice that MS analysis of the crude did not show the presence of dimers or higher oligomers.

Assessment of the structure of the isolated compound was carried out taking advantage of the differential reactivity of thiols and thioethers with H_2O_2, which oxidizes thiols to sulfonic acids and thioethers to sulfoxides (Scheme S2). MS analysis after treatment with H_2O_2 showed that reaction with **3a** afforded the sulfonic acid (mass increment: 48), while reaction with **5a** provided the sulfoxide (mass increment: 16). These results confirmed that the isolated compound was the cyclic oligonucleotide (**5a**), and showed that under the retro-Diels–Alder conditions maleimide deprotection was accompanied by cyclization.

The whole process was repeated (Scheme 1a, route B), with the only difference that maleimide deprotection was carried out in a microwave oven using conditions tested in previous experiments (12a). Reproducibility was verified, and in this case the target **5a** was found to be present in the crude in even higher yield (80-85 %, HPLC). Hence, use of microwave irradiation to promote the retro-Diels–Alder reaction resulted in optimization of the cyclization methodology.

It is tempting to hypothesize that intramolecular cyclization might be due to oligonucleotide **4a** forming a hairpin-type structure rather than a duplex-type one (Figure S4), which might facilitate formation of dimers, but any of these structures would hardly survive in the conditions that deprotect the maleimide. In order to get more insight into the scope of the method, additional experiments were carried out with oligonucleotides **1b**, **1c** and **1d**. Sequences **b** (dT_{10}) and **c** cannot preorganize or fold as a hairpin and favor cyclization over oligomerization. End-modified, CpG-containing 5-mers such as oligonucleotide **c** ($^{5'}$dCGTTA) have been described to induce apoptosis in leukaemia cell lines (16). As to oligonucleotide **d**, the sequence of this 26-mer is that of the trinucleotide repeat present in the 3' UTR region of the dystrophia myotonica protein kinase gene (*DMPK*, located on chromosome 19), whose expansion leads to the development of DM1 (myotonic dystrophy type 1) (17).

Maleimido- and thiol-protected precursors **1b**, **1c** and **1d** (Scheme 1a) were synthesized as described for **1a**, and submitted to the optimized thiol deprotection/purification/retro-Diels–Alder reaction series of steps. Again, MS analysis confirmed the structure of purified intermediates **3b**, **3c** and **3d**, and that the products formed after retro Diels-Alder reaction (**5b**, **5c**, **5d**) were the cyclic oligonucleotides (there was an increase of 16 mass units after reaction with H_2O_2). These results corroborate the usefulness of this methodology to prepare cyclic oligonucleotides with a 3-thiosuccinimide linkage. It is also worth noticing that cyclization yields (68-89 %) were either of the same order as that those reported in the literature or superior.

This work has provided various pieces of evidence. First, that TCEP does react with free maleimides, furnishing intermediates that finally evolve into succinimides. We have also found that oligonucleotides with masked maleimides and thiols at either end (**1**) can be cyclized following route B (thiol deprotection followed by maleimide deprotection) provided that the excess TCEP used for disulfide reduction is separated from the [protected maleimido]-oligonucleotide-SH (**3**). Importantly, maleimide deprotection and cyclization take place at the same step, and in the high dilution reaction conditions (25 μM) the cyclic compound is the main one in the crude, as assessed by HPLC. This methodology has proved useful to obtain small- to medium-size cyclic oligonucleotides with excellent yield and purity regardless of sequence and length, in particular if maleimide deprotection and cyclization are carried out in a microwave oven. Intramolecular cyclization takes place under conditions (90 ^{o}C) hardly compatible with internal (or external) associations of the chain, which precludes internal preorganization or use of an external template to bring the ends of the chain close to each other. Therefore, formation of the target cyclic oligonucleotide seems to be driven essentially by the high dilution of the reaction medium and the high affinity of thiols for maleimides.

With very few exceptions (9a, 11a), existing methods not relying on structural preorganization have not been able to provide cyclic oligonucleotides with more than 20 nucleotides, and cyclization yields are often moderate. Commercial synthons exist for the

preparation of disulfide-bridged cyclic oligonucleotides (9a), but when preparing medium size circles (*ca* 20-mers) by thiol to disulfide oxidation, we have often obtained mixtures of monomers and dimers even working at oligonucleotide concentrations much lower than those recommended (18). Derivatives for the synthesis of cyclic oligonucleotides using the Huisgen cycloaddition are also commercially available. This synthetic alternative has furnished the longest circle reported (72-mer), but the cyclization yield was not provided (11a).

The methodology here described allows for the straightforward preparation of cyclic oligonucleotides. We believe it is particularly well-suited to prepare medium size circles that would be difficult to obtain with all phosphodiester linkages by any other approach, and which are conformationally-constrained models of non B-DNA structures (**5d** is an example). All the necessary synthons are commercially available, and no additional reagents such as a metal catalyst are required for the cyclization. This is not without importance, because the oligonucleotide may be degraded if accompanied by Cu(I) ions (8f).

We do not know what is the longest oligonucleotide that can be cyclized using this chemistry, but probably much longer than the 26-mer here reported. The answer to this question requires different procedures to verify that the cyclic molecule has been obtained (7c), since the mass spectrometric characterization described in this paper is no longer easily feasible. Work is in progress to further assess the scope of this methodology.

Acknowledgements. This work was supported by funds from the Ministerio de Ciencia e Innovación (grant CTQ2010-21567-C02-01, and the project RNAREG, grant CSD2009-00080, funded under the programme CONSOLIDER INGENIO 2010), and the Generalitat de Catalunya (2009SGR-208). A. S. was a recipient fellow of the Generalitat de Catalunya.

Notes and references

[a] *Departament de Química Orgànica and IBUB, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain. Fax: +34 933397878; Tel: + 34 934021263; E-mail: anna.grandas@ub.edu*

† Electronic Supplementary Information (ESI) available: Synthesis and characterization (HPLC, MALDI-TOF and ESI MS) of compounds **1**, **3** and **5**; characterization of oxidized **3** and **5** and of side products. See DOI: 10.1039/b000000x/

1 *(a)* E. T. Kool, *Annu. Rev. Biophys. Biomol. Struct.*, 1996, **25**, 1-28; *(b)* E. T. Kool, *Acc. Chem. Res.*, 1998, **31**, 502-510.

2 *(a)* D. Liu, S. L. Daubendiek, M. A. Zillmann, K. Ryan and E. T. Kool, *J. Am. Chem. Soc.,* 1996, **118**, 1587-1594; *(b)* M. Frieden, E. Pedroso and E. T. Kool, *Angew. Chem. Int. Edit.*, 1999, **38**, 3654-3657.

3 *(a)* S. A. Salisbury, S. E. Wilson, H. R. Powell, O. Kennard, P. Lubini, G. M. Sheldrick, N. Escaja, E. Alazzouzi, A. Grandas and E. Pedroso, *Proc. Natl. Acad. Sci. U.S.A.,* 1997, **94**, 5515-5518; *(b)* N. Escaja, E. Pedroso, M. Rico and C. González, *J. Am. Chem. Soc.,* 2000, **122**, 12732-12742; *(c)* N. Escaja, J. L. Gelpí, M. Orozco, M. Rico, E. Pedroso and C. González, *J. Am. Chem. Soc.,* 2003, **125**, 5654-52; *(d)* N. Escaja, I. Gómez-Pinto, E. Pedroso and C. González, *J. Am. Chem. Soc.,* 2007, **129**, 2004-2014.

4 *(a)* M. Hyodo and Y. Hayakawa, *Bull. Chem. Soc. Jpn.*, 2004, **77**, 2089-2093; *(b)* G. Witte, S. Hartung, K. Büttner and K.-P. Hopfner, *Mol. Cell*, 2008, *30*, 167-178; *(c)* C. Solano, B. García, C. Latasa, A. Toledo-Arana, V. Zorraquino, J. Valle, J. Casals, E. Pedroso and I. Lasa, *Proc. Nat. Acad. Sci. USA*, 2009, **106**, 7007-8002; *(d)* R. Hengge, *Nat. Rev. Microbiol.*, 2009, *7* 263-273; *(e)* A. Grajkowski, J. Cieslak, A. Gapeev, C. Schindler and S. L. Beaucage, *Bioconjugate Chem.*, 2010, **21**, 2147-2152; *(f)* J. J. Woodward, A. T. Iavarone and D. A. Portnoy, *Science*, 2010, *328*, 1703-1705.

5 *(a)* L. De Napoli, A. Galeone, L. Mayol, A. Messere, D. Montesarchio and G. Piccialli, *Bioorg. Med. Chem.*, 1995, **3**, 1325-1329; *(b)* E. Alazzouzi, N. Escaja, A. Grandas and E. Pedroso, *Angew. Chem. Int. Edit.*, 1997, **36**, 1506-1508; *(c)* R. Micura, *Chem. Eur. J.*, 1999, **5**, 2077-2082.

6 M. Smietana and E. T. Kool, *Angew. Chem. Int. Edit.*, 2002, **41**, 3704-3707.

7 *(a)* G. Prakash and E. T. Kool, *Chem. Commun.*, 1991, 1161-1163; *(b)* A. M. Diegelman and E. T. Kool, in *Current Protocols in Nucleic Acid Chemistry* (Chapter 5: Methods for Cross-Linking Nucleic Acids), eds. M. Egli, P. Herdewijn, A. Matsuda and Y. S. Sanghvi, John Wiley and Sons, 2000, pp. 5.2.1-5.2.27; *(c)* M. G. Grimau, A. Aviñó, R. Gargallo and R. Eritja, *Chem. Biodivers.*, 2005, **2**, 275-285.

8 *(a)* G. W. Ashley and D. M. Kushlan, *Biochemistry* 1991, **30**, 2927-2933; *(b)* J. Chen, D. Liu, A. H. F. Lee, J. Qi, A. S. C. Chan and T. Li, *Chem. Commun.*, 2002, 2686-2687; *(c)* D. Liu, J. Chen, A. H. F. Lee, L. M. C. Chow, A. S. C. Chan and T. Li, *Angew. Chem. Int. Edit.*, 2003, **42**, 797-799; *(d)* D. A Di Giusto, W. A. Wlassoff, J. J. Gooding, B. A. Messerle and G. C. King, *Nucleic Acids Res.*, 2005, **33**, e64; *(e)* J. Qi and R. H. Shafer, *Nucleic Acids Res.*, 2005, **33**, 3185-3192; *(f)* A. H. El-Sagheer and T. Brown, in *Current Protocols in Nucleic Acids Chemistry* (Chapter 4: Synthesis of Modified Oligonucleotides and Conjugates), eds. M. Egli, P. Herdewijn, A. Matsuda and Y. S. Sanghvi, John Wiley and Sons, 2008, pp. 4.33.1-4.33.21; *(g)* T. Zhou, X. Li, M. T. T. Ng, Y. Wang, N. M. Quek, J. Luo, W. Yuan, C.-H. Tan, H. Zen and T. Li, *Bioconjugate Chem.*, 2009, **20**, 644-647.

9 *(a)* H. Gao, M. Yang, R. Patel and A. F. Cook, *Nucleic Acids Res.* 1995, **23**, 2025-2029; *(b)* O. P. Edupuganti, E. Defrancq and P. Dumy, *J. Org. Chem.*, 2003, **68**, 8708-8710.

10 H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2010, **40**, 2004-2012.

11 *(a)* R. Kumar, A. El-Sagheer, J. Tumpane, P. Lincoln, L. M. Wilhelmsson and T. Brown, *J. Am. Chem. Soc.*, 2007, **129**, 6859-6864; *(b)* A. El-Sagheer, R. Kumar, S. Findlow, J. M. Werner, A. N. Lane and T. Brown, *ChemBioChem*, 2008, *9*, 50-52; *(c)* J. Lietard, A. Meyer, J.-J. Vasseur and F. Morvan, *J. Org. Chem.*, 2008, **73**, 191-200; *(d)* M. Nakane, S. Ichikawa and A. Matsuda, *J. Org. Chem.*, 2008, **73**, 1842-1851; *(e)* G. Pourceau, A. Meyer, J.-J. Vasseur and F. Morvan, *J. Org. Chem.*, 2009, **74**, 6837-6842.

12 *(a)* A. Sánchez, E. Pedroso and A. Grandas, *Org. Lett.*, 2011, **13**, 4363-4367; *(b)* A. Sánchez, E. Pedroso and A. Grandas, *Bioconjugate Chem.*, 2012, **23**, 300-307.

13 E. B. Getz, M. Xiao, T, Chakrabarty, R. Cooke and P. R. Selvin, *Anal. Biochem.* 1999, **273**, 73-80.

14 M. Koszytkowska-Stawinska and W. Sas, *J. Chem. Res. (S)*, 1999, 694-695.

15 B. Pal, P. K. Pradhan, P. Jaisankar and V. S. Gir, *Synthesis*, 2003, 1549-1552.

16 *(a)* D. M. Tidd, D. G. Spiller, C. M. Broughton, L. C. Norbury, R. E. Clark and R. V. Giles, *Nucleic Acids Res.*, 2000, **28**, 2242-2250; *(b)* D. M. Tidd, C. M. Broughton and R. E. Clark, *Nucleic Acids Res.*, 2003, **31**, 2297-2304.

17 *(a)* J. P. Orengo, P. Chambon, D. Metzger, D. R. Mossier, G. J. Snipes and T. A. Cooper, *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 2646-2651; *(b)* J. D. Cleary, S. Tomé, A. López Castel, G. B. Panigrahi, L. Foiry, K. A. Hagerman, H. Sroka, D. Chitayat. G. Gourdon and C. E. Pearson, 2010, *Nat. Struct. Mol Biol.*, 2010, **17**, 1079-1087.

18 N. Escaja and E. Pedroso, unpublished results.