Electronic Supplementary Material (ESI) for Chemical Communications

SUPPORTING INFORMATION TO:

Oligonucleotide cyclization: The thiol-maleimide reaction revisited.

Albert Sánchez, Enrique Pedroso, Anna Grandas*

1. General materials and methods.

Nucleoside phosphoramidites (dA^{Pac}, dC^{Ac}, dG^{iPrPac} or dG^{Dmf} and dT), CPG support to obtain $^{3'}$thiol-modified oligonucleotides (3'-Thiol-Modifier C3 S-S CPG = 1-*O*-(4,4'-Dimethoxytrityl)-propyl-disulfide, 1'-succinyl-LCAA CPG support; A. Kumar, S. Advani, H. Dawar, G. P. Talwar, *Nucleic Acids Res.* **1991**, *19*, 4561), and oligonucleotide synthesis reagents were from either Link Technologies, Glen Research or Applied Biosystems. [Protected maleimido]-phosphoramidite (see structure in Scheme S1) was synthesized as previously described (A. Sánchez, E. Pedroso, A. Grandas, *Org. Lett.* **2011**, *13*, 4364-4367), but is available from Glen Research.

Microwave irradiation was carried out in a Biotage Initiator^{TM} oven, using reaction vessels of 0.5-2 mL. Samples were lyophilized in a FreezeMobile Virtis instrument.

TCEP·HCl was purchased from Aldrich and 35 % (w/w) aq. hydrogen peroxide solution was from Alfa Aesar.

HPLC.

Reversed-phase HPLC analysis and purification was performed using analytical and semipreparative Waters or Shimadzu systems. Unless otherwise indicated, analysis and purification conditions were:

Oligonucleotide analysis conditions: Kromasil C18 column (10 μm, 100 Å, 250 × 4.0 mm) from Akzo Nobel; solvent A: 0.1 M triethylammonium acetate, solvent B: H_2O/ACN 1:1 (v/v), flow: 1 mL/min, detection wavelength: 254 nm.

Oligonucleotide purification conditions (semipreparative scale): Jupiter C18 column (10 μm, 300 Å, 250 × 10.0 mm) from Phenomenex; solvent A: 0.1 M triethylammonium acetate, solvent B: H_2O/ACN 1:1 (v/v), flow: 3 mL/min detection wavelength: 260 nm.

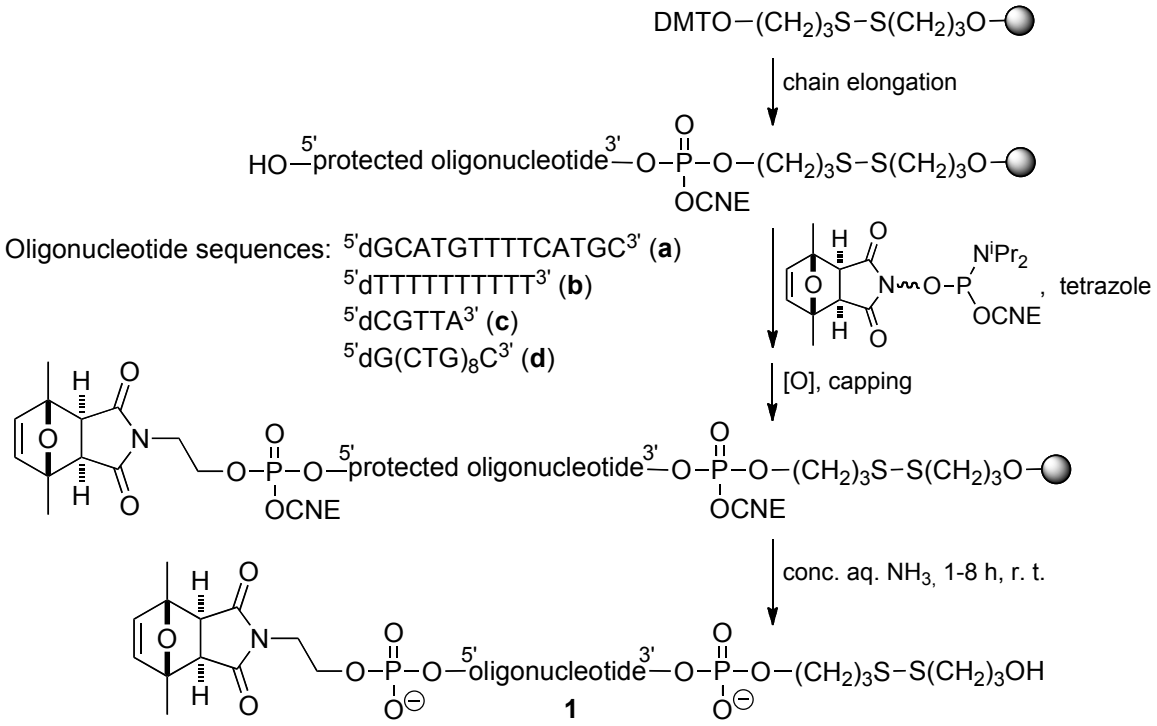
Mass spectrometry.

MALDI-TOF mass spectra were recorded on a 4800 *Plus* ABSciex instrument. Unless indicated reflector mode was used for the analysis of the products. Typical oligonucleotide analysis conditions: 1:1 (v/v) 2,4,6-trihydroxyacetophenone/ammonium citrate (THAP/CA), negative mode. ESI mass spectra were obtained using an LC/MSD-TOF spectrometer from Agilent Technologies.

Abbreviations.

ACN=acetonitrile, CNE=2-cyanoethyl, Dmf=dimethylaminomethylene (dimethylformamidine), DMT=4,4'-dimethoxytrityl, iPrPac=isopropylphenoxyacetyl, LCAA-CPG=long chain aminoalkyl controlled pore glass beads, Pac=phenoxyacetyl, TCA=trichloroacetic acid, TCEP =tris(2-carboxyethyl)phosphine.

2. Solid phase synthesis 5'-[protected maleimido]-3'-[protected thiol]-oligonucleotides (1).



Scheme S1. Synthesis of oligonucleotides **1** (linear precursors of cyclic oligonucleotides).

Oligonucleotide chains were assembled in a 3400 ABI automatic synthesizer at the 1 μmol scale using the 3'-Thiol-Modifier CPG support (Scheme S1). Oligonucleotide elongation was carried out using the phosphite triester approach (standard synthesis cycles) and a 0.02 M iodine solution in the oxidation step to avoid oxidative cleavage of the disulfide linkage to the support. [Protected maleimido]-phosphoramidite was incorporated at the 5' end as any standard nucleoside-phosphoramidite. After chain elongation, treatment with concd. aq. ammonia at room temperature removed protecting groups from the oligonucleotide (1 h in case of homo-dT oligomers; 4 or 8 h when all the bases were present depending on whether G was protected with the iPrPac or Dmf groups, respectively). After filtration and washing, ammonia was evaporated under reduced pressure, and the sample lyophilized. Unless the homogeneity of the crudes was sufficiently high (as assessed by HPLC), oligonucleotides **1** were purified by reversed-phase HPLC (Figure S1). All oligonucleotides were quantified by UV spectroscopy (λ = 260 nm), and characterized by MALDI-TOF or ESI-TOF mass spectrometry.

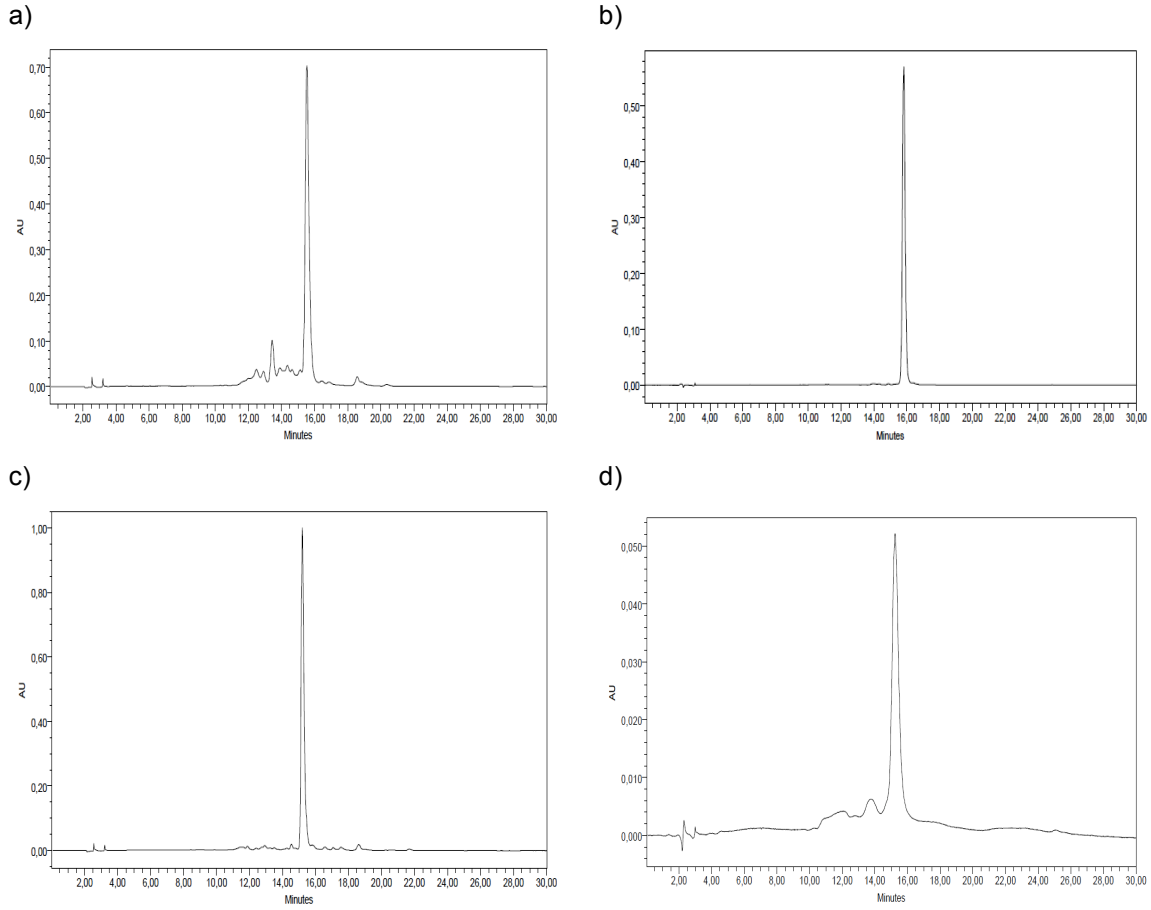


Figure S1. Crude oligonucleotides **1**. a) [Protected maleimido]-$^{5'}$dGCATGTTTTCATGC$^{3'}$-$(CH_2)_3$S-S$(CH_2)_3$OH (**1a**); b) [protected maleimido]-$^{5'}$dTTTTTTTTTT$^{3'}$-$(CH_2)_3$S-S$(CH_2)_3$OH (**1b**); c) [protected maleimido]-$^{5'}$dCGTTA$^{3'}$-$(CH_2)_3$S-S$(CH_2)_3$OH (**1c**); d) [protected maleimido]-$^{5'}$dG$(CTG)_8$C$^{3'}$-$(CH_2)_3$S-S$(CH_2)_3$OH (**1d**).

*[Protected maleimido]-$^{5'}$dGCATGTTTTCATGC$^{3'}$-$(CH_2)_3$S-S$(CH_2)_3$OH (**1a**):*

Analytical HPLC (Figure S1a, gradient from 10 to 70 % of B in 30 min): t_R = 15.5 min (82 % in the crude); MALDI-TOF MS (negative mode): *m/z* 4784.1 $[M-H]^-$, M calcd. for $C_{155}H_{202}N_{47}O_{96}P_{15}S_2$ 4785.8.

*[Protected maleimido]-$^{5'}$dTTTTTTTTTT$^{3'}$-$(CH_2)_3$S-S$(CH_2)_3$OH (**1b**):*

Analytical HPLC (Figure S1b, gradient from 10 to 70 % of B in 30 min): t_R = 15.8 min (> 99 % in the crude); MALDI-TOF MS (negative mode): *m/z* 3520.4 $[M-H]^-$, M calcd. for $C_{118}H_{158}N_{21}O_{78}P_{11}S_2$ 3521.6.

*[Protected maleimido]-$^{5'}$dCGTTA$^{3'}$-$(CH_2)_3$S-S$(CH_2)_3$OH (**1c**):*

Analytical HPLC (Figure S1c, gradient from 10 to 70 % of B in 30 min): t_R = 15.2 min (96 % in the crude); MALDI-TOF MS (negative mode): *m/z* 2019.1 $[M-H]^-$, M calcd. for $C_{67}H_{90}N_{18}O_{39}P_6S_2$ 2020.4.

*[Protected maleimido]-$^{5'}$dG$(CTG)_8$C$^{3'}$-$(CH_2)_3$S-S$(CH_2)_3$OH (**1d**):*

Analytical HPLC (Figure S1d, gradient from 10 to 70 % of B in 30 min): t_R = 15.2 min (77 % in the crude); ESI-TOF MS (negative mode): *m/z* 1412.5 $[M-6H]^{6-}$, 1210.6 $[M-7H]^{7-}$, 1059.1 $[M-8H]^{8-}$, 941.2 $[M-9H]^{9-}$, 847.0 $[M-10H]^{10-}$, 769.9 $[M-11H]^{11-}$, 705.7 $[M-12H]^{12-}$, 651.3 $[M-13H]^{13-}$, 604.7 $[M-14H]^{14-}$, 564.3 $[M-15H]^{15-}$, 529.0 $[M-16H]^{16-}$, 497.8 $[M-17H]^{17-}$, M found 8480.2, M calcd. for $C_{269}H_{348}N_{89}O_{172}P_{27}S_2$ 8476.4.

3. Reduction of the disulfide linkage of 1 using TCEP.

TCEP·HCl (10 equiv) was added to the [protected maleimido]-$^{5'}$oligonucleotide$^{3'}$-$(CH_2)_3S$-$S(CH_2)_3OH$ **1** (0.5 mM aq. solution). The required amount of a 5 % aq. NaOH solution was added to increase the initial pH value of the solution (pH ≈ 2-3) to pH = 4.9, and the mixture was stirred for 1 h at room temperature. Reaction crudes were analyzed by HPLC and the resulting [protected maleimido]-$^{5'}$oligonucleotide$^{3'}$-SH compounds (**3**) were purified by HPLC and characterized by MALDI-TOF or ESI-TOF MS.

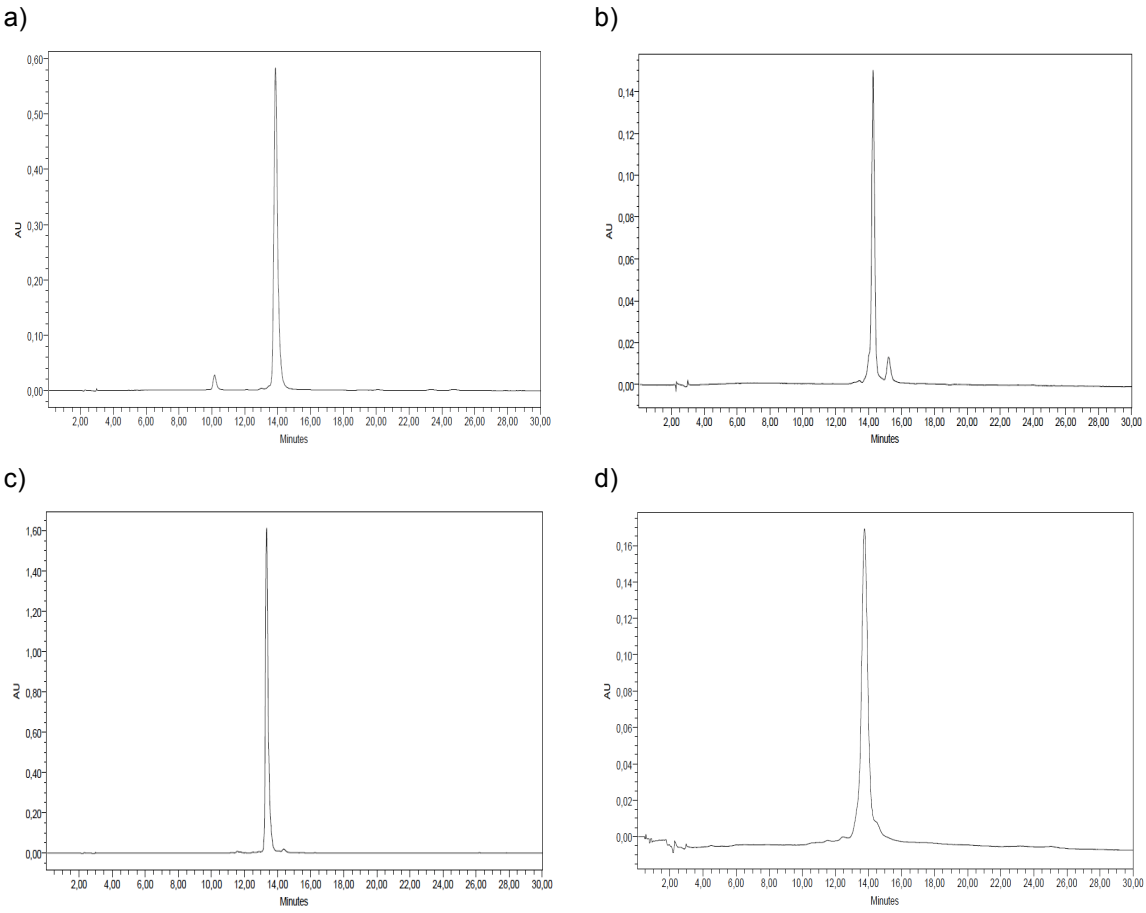


Figure S2. Crude oligonucleotides **3** after disulfide reduction using TCEP. a) [Protected maleimido]-$^{5'}$dGCATGTTTTCATGC$^{3'}$-$(CH_2)_3SH$ (**3a**); b) [protected maleimido]-$^{5'}$dTTTTTTTTTT$^{3'}$-$(CH_2)_3SH$ (**3b**); c) [protected maleimido]-$^{5'}$dCGTTA$^{3'}$-$(CH_2)_3SH$ (**3c**); d) [protected maleimido]-$^{5'}$dG$(CTG)_8$C$^{3'}$-$(CH_2)_3SH$ (**3d**).

*[Protected maleimido]-$^{5'}$dGCATGTTTTCATGC$^{3'}$-$(CH_2)_3$SH (**3a**):*

Analytical HPLC (Figure S2a, gradient from 10 to 70 % of B in 30 min): t_R = 13.9 min (97 % in the crude); MALDI-TOF MS (negative mode): *m/z* 4695.6 $[M-H]^-$, M calcd. for $C_{152}H_{196}N_{47}O_{95}P_{15}S$ 4695.8.

*[Protected maleimido]-$^{5'}$dTTTTTTTTTT$^{3'}$-$(CH_2)_3$SH (**3b**):*

Analytical HPLC (Figure S2b, gradient from 10 to 70 % of B in 30 min): t_R = 14.3 min (94 % in the crude); MALDI-TOF MS (negative mode): *m/z* 3430.1 $[M-H]^-$, M calcd. for $C_{115}H_{152}N_{21}O_{77}P_{11}S$ 3431.5.

*[Protected maleimido]-$^{5'}$dCGTTA$^{3'}$-$(CH_2)_3$SH (**3c**):*

Analytical HPLC (Figure S2c, gradient from 10 to 70 % of B in 30 min): t_R = 13.3 min (> 99 % in the crude); MALDI-TOF MS (negative mode): *m/z* 1929.6 $[M-H]^-$, M calcd. for $C_{64}H_{84}N_{18}O_{38}P_6S$ 1930.3.

*[Protected maleimido]-$^{5'}$dG$(CTG)_8$C$^{3'}$-$(CH_2)_3$SH (**3d**):*

Analytical HPLC (Figure S2d, gradient from 10 to 70 % of B in 30 min): t_R = 13.8 min (99 % in the crude); ESI-TOF MS (negative mode): *m/z* 1396.7 $[M-6H]^{6-}$, 1197.0 $[M-7H]^{7-}$, 1047.3 $[M-8H]^{8-}$, 930.8 $[M-9H]^{9-}$, 837.6 $[M-10H]^{10-}$, 761.4 $[M-11H]^{11-}$, 697.8 $[M-12H]^{12-}$, 644.1 $[M-13H]^{13-}$, 598.0 $[M-14H]^{14-}$, 558.1 $[M-15H]^{15-}$, M found 8386.2, M calcd. for $C_{266}H_{342}N_{89}O_{171}P_{27}S$ 8386.3.

4. Microwave-promoted one-pot maleimide deprotection and cyclization of 5'-[protected maleimido]-3'-thiol-oligonucleotides 3.

A solution (1000 µL) of [protected maleimido]-$^{5'}$oligonucleotide$^{3'}$-SH **3** in a 1:1 (v/v) MeOH/H_2O mixture (25 µM concentration) was introduced in a microwave vial and irradiated for 90 min at 90 ^{o}C. Solvent was removed under vacuum, and the resulting crude was dissolved in water for HPLC analysis. The main product was collected and characterized by MALDI-TOF or ESI-TOF MS.

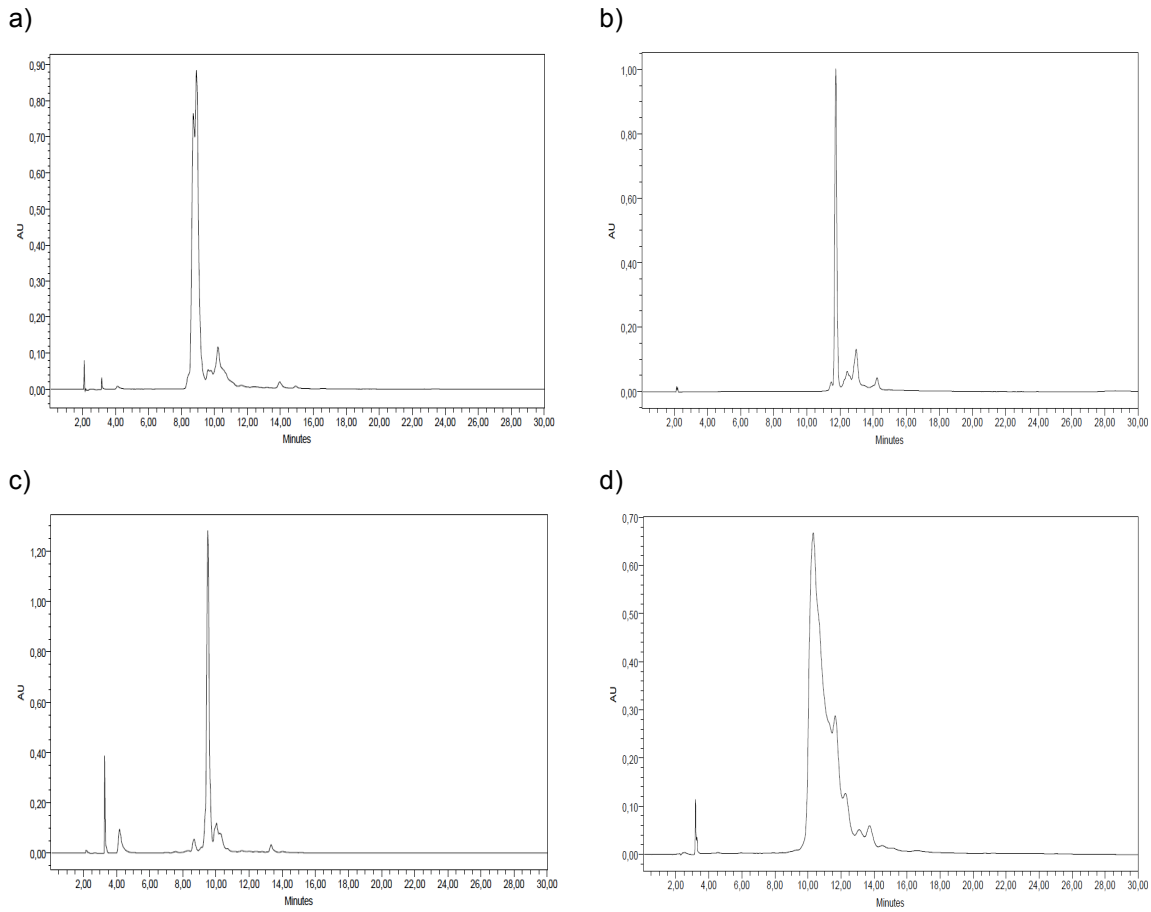


Figure S3. Crude cyclic oligonucleotides **5**. a) Cyclic dGCATGTTTTCATGC (**5a**); b) cyclic dTTTTTTTTTT (**5b**); c) cyclic dCGTTA (**5c**); d) cyclic dG$(CTG)_8$C (**3d**).

*Cyclic dGCATGTTTTCATGC (**5a**):*

Analytical HPLC (Figure S3a, gradient from 10 to 70 % of B in 30 min): t_R = 8.7 and 8.9 min (two diastereomers, 89 % in the crude); MALDI-TOF MS (negative mode): *m/z* 4599.5 $[M-H]^-$, M calcd. for $C_{146}H_{188}N_{47}O_{94}P_{15}S$ 4599.7.

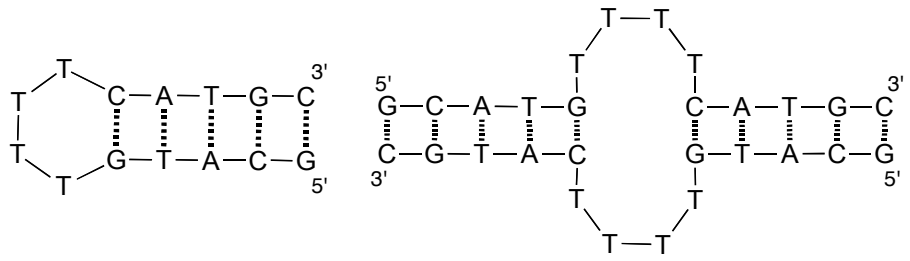


Figure S4. Possible associations (intra- and intermolecular) for oligonucleotide sequence **a**. Any of these structures would hardly survive in the conditions that deprotect the maleimide and promote cyclization.

*Cyclic dTTTTTTTTTT (**5b**):*

Analytical HPLC (Figure S3b, gradient from 10 to 70 % of B in 30 min): t_R = 11.8 min (71 % in the crude); MALDI-TOF MS (negative mode): *m/z* 3334.4 $[M-H]^-$, M calcd. for $C_{109}H_{144}N_{21}O_{76}P_{11}S$ 3335.5.

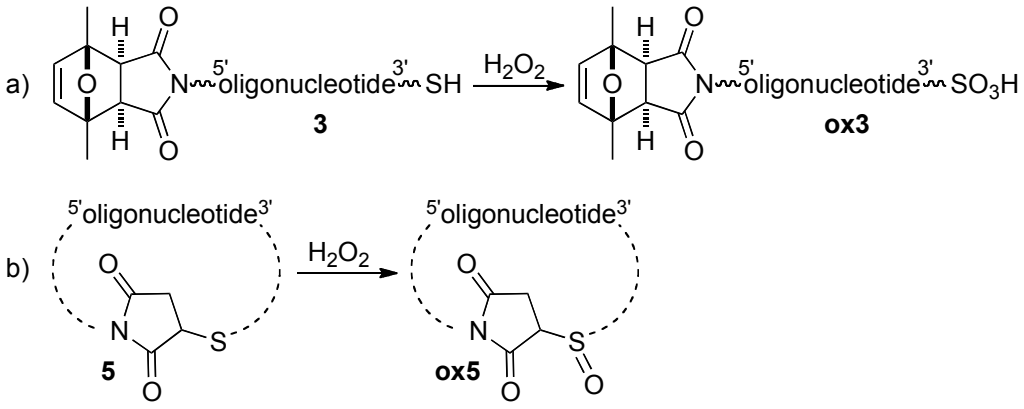
*Cyclic dCGTTA (**5c**):*

Analytical HPLC (Figure S3c, gradient from 10 to 70 % of B in 30 min): t_R = 9.5 min (71 % in the crude); MALDI-TOF MS (negative mode): *m/z* 1833.6 $[M-H]^-$, M calcd. for $C_{58}H_{76}N_{18}O_{37}P_6S$ 1834.3.

*Cyclic dG(CTG)$_8$C (**5d**):*

Analytical HPLC (Figure S3d, gradient from 10 to 70 % of B in 30 min): t_R = 10.3 min (two diastereomers, 68 % in the crude); ESI-TOF MS (negative mode): *m/z* 1380.7 $[M-6H]^{6-}$, 1183.3 $[M-7H]^{7-}$, 1035.3 $[M-8H]^{8-}$, 920.1 $[M-9H]^{9-}$, 828.1 $[M-10H]^{10-}$, 752.6 $[M-11H]^{11-}$, 689.8 $[M-12H]^{12-}$, 636.7 $[M-13H]^{13-}$, 591.3 $[M-14H]^{14-}$, 551.7 $[M-15H]^{15-}$, 517.1 $[M-16H]^{16-}$, 486.6 $[M-17H]^{17-}$, M found 8290.2, M calcd. for $C_{260}H_{334}N_{89}O_{170}P_{27}S$ 8290.3.

5. Assessment of circularity: Reaction with H_2O_2.



Scheme S2. Assessment of cyclization: Reaction with H_2O_2 oxidizes thiol-containing linear precursors such as **3** to sulfonic acids **ox3** (a), while the thioethers formed after cyclization **5** are oxidized to sulfoxide **ox5** (b).

2 μL of a 3.5 % (w/w) aq.H_2O_2 solution was added to 2 μL of an aq. solution of oligonucleotide (0.25 mM). The mixture was left to react for 1 h at room temperature, after which time it was analyzed by MALDI-TOF or ESI-TOF MS. This procedure was used to analyze both [protected maleimido]-$^{5'}$oligonucleotide$^{3'}$-SH (**3**) and cyclic oligonucleotides (**5**), and furnished the sulfonic acid (**ox3**) and thioether (**ox5**), respectively (Scheme S2).

*Cyclic dGCATGTTTTCATGC (**5a**) + H_2O_2:*
MALDI-TOF MS (negative mode): *m/z* 4615.1 $[M-H]^-$, M calcd. for $C_{146}H_{188}N_{47}O_{95}P_{15}S$ 4615.7.

*Cyclic dTTTTTTTTTT (**5b**) + H_2O_2:*
MALDI-TOF MS (negative mode): *m/z* 3350.4 $[M-H]^-$, M calcd. for $C_{109}H_{144}N_{21}O_{77}P_{11}S$ 3351.5

*Cyclic dCGTTA (**5c**) + H_2O_2:*
MALDI-TOF MS (negative mode): *m/z* 1849.5 $[M-H]^-$, M calcd. for $C_{58}H_{76}N_{18}O_{38}P_6S$ 1850.3.

*Cyclic dG(CTG)$_8$ (**5d**) + H_2O_2:*
ESI-TOF MS (negative mode): *m/z* 1383.4 $[M-6H]^{6-}$, 1185.6 $[M-7H]^{7-}$, 1037.3 $[M-8H]^{8-}$, 921.9 $[M-9H]^{9-}$, 829.6 $[M-10H]^{10-}$, 754.1 $[M-11H]^{11-}$, 691.2 $[M-12H]^{12-}$, 637.9 $[M-13H]^{13-}$, 592.3 $[M-14H]^{14-}$, 552.8 $[M-15H]^{15-}$, 518.1 $[M-16H]^{16-}$, 487.6 $[M-17H]^{17-}$, M found 8306.2, M calcd. for $C_{260}H_{334}N_{89}O_{171}P_{27}S$ 8306.3.

*[Protected maleimido]-$^{5'}$dGCATGTTTTCATGC$^{3'}$-$(CH_2)_3$SH (**3a**) + H_2O_2:*

MALDI-TOF MS (negative mode): *m/z* 4743.6 (linear) $[M-H]^-$, M calcd. for $C_{152}H_{196}N_{47}O_{98}P_{15}S$ 4743.8.

*[Protected maleimido]-$^{5'}$dTTTTTTTTTT$^{3'}$-$(CH_2)_3$SH (**3b**) + H_2O_2:*

MALDI-TOF MS (negative mode): *m/z* 3478.3 $[M-H]^-$, M calcd. for $C_{115}H_{152}N_{21}O_{80}P_{11}S$ 3479.5.

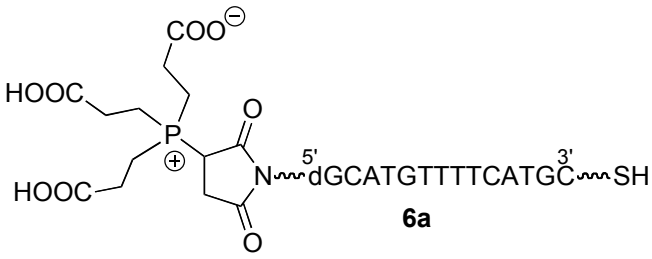
*[Protected maleimido]-$^{5'}$dCGTTA$^{3'}$-$(CH_2)_3$SH (**3c**) + H_2O_2:*

MALDI-TOF MS (negative mode): *m/z* 1977.5 $[M-H]^-$, M calcd. for $C_{64}H_{84}N_{18}O_{41}P_6S$ 1978.3.

*[Protected maleimido]-$^{5'}$dG(CTG)$_8$C$^{3'}$-$(CH_2)_3$SH (**3d**) + H_2O_2:*

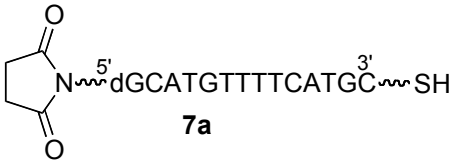
ESI-TOF MS (negative mode): *m/z* 842.4 $[M-10H]^{10-}$, 765.8 $[M-11H]^{11-}$, 701.9 $[M-12H]^{12-}$, 647.8 $[M-13H]^{13-}$, 601.4 $[M-14H]^{14-}$, 561.3 $[M-15H]^{15-}$, 526.1 $[M-16H]^{16-}$, 495.1 $[M-17H]^{17-}$, M found 8434.2, M calcd. for $C_{266}H_{342}N_{89}O_{174}P_{27}S$ 8434.3

6. Characterization of side products.



6a

Analytical HPLC (gradient from 10 to 70 % of B in 30 min): t_R = 9.6 min; MALDI-TOF MS (negative mode): *m/z* 4848.4 $[M-H]^-$, M calcd. for $C_{155}H_{203}N_{47}O_{100}P_{16}S$ 4849.8.



7a

Analytical HPLC (gradient from 10 to 70 % of B in 30 min): t_R = 11.5 min; MALDI-TOF MS (negative mode): *m/z* 4600.5 $[M-H]^-$, M calcd. for $C_{146}H_{190}N_{47}O_{94}P_{15}S$ 4601.7.