Tetrahedron 69 (2013) 7847–7853

Contents lists available at SciVerse ScienceDirect

Tetrahedron

journal homepage: www.elsevier.com/locate/tet

Toward encoding reactivity using double-stranded DNA. Sequence-dependent native chemical ligation of DNA binding polyamides

Adrian Jiménez-Balsa, Verónica I. Dodero, José L. Mascareñas *

Departamento de Química Orgánica and Centro Singular de Investigación en Química Biolóxica e Materiales Moleculares (CIQUS). Universidade de Santiago de Compostela, 15782, Santiago de Compostela, Spain

ARTICLE INFO

Article history:
Received 7 March 2013
Received in revised form 21 May 2013
Accepted 28 May 2013
Available online 4 June 2013

Keywords:
DNA binding
Native chemical ligation
Molecular recognition
Polyamides

ABSTRACT

The coupling reaction between distamycin-related polyamides equipped with cysteine and thioester groups can be accelerated in the presence of double-stranded (ds) oligonucleotides with selected sequences. While the coupling reaction of reactive partners containing three pyrrole units is accelerated by dsDNAs containing ATTTTA or ATGTTA sites, the heterodimeric coupling between a tripyrrole and a polyamide equipped with two pyrroles and one imidazole, is accelerated by the latter DNA, but not by the other containing the A/T rich tract. These differences can be exploited for selecting preferred coupling pairs from a mixture of reactive monomers; therefore the reaction outcome depends on the instructions provided by the dsDNA sequence.

1. Introduction

The processes of gene transcription and replication rely on the DNA-templated polymerization of activated nucleotide units. The template effect increases the effective molarity of the reactants, therefore favoring the coupling reactions, and ensures that the biopolymeric product exhibits the correct sequence.[1] Inspired by these natural systems, chemists have long pursued the use of DNA templates to promote and control synthetic transformations of non-biological reactants. Although the first examples on this topic were published in the 1990s,[2] it was not until recently that DNA-template synthesis (DTS) has emerged as a reliable and powerful tool.[3] The applications of this technology are not restricted to the field of chemical synthesis, but also reach other areas, such as reaction discovery, drug development, and biosensing.[4] Despite all the advances, the methodology still suffers from some limitations, such as the requirement of tethering the reactants to linear DNA chains, which are usually not desired parts of the final products. On the other hand, achieving catalytic turnover is certainly difficult owing to the usually higher affinity of the products for the DNA template.[5]

Natural DNA-encoded processes are not restricted to those controlled by Watson–Crick base pairings in single stranded DNA (ssDNA) chains. Other extremely relevant cellular events, like the regulation of gene transcription, also depend on the correct reading of specific DNA sequences.[6] However, in this case, the information is encoded in the surface of the major and minor grooves in the double-stranded DNA (dsDNA). Although nature does not use dsDNAs for templating reactivity, the implicit recognition code provided by dsDNA, together with its unique structural features, suggest that it might offer an interesting potential for promoting chemical transformations in a programmed manner. Up to now, the use of double-stranded DNA as an instructed reactivity template has been scarcely explored,[7] and is essentially restricted to the examples reported by Dervan's and co-workers on a triplex helix-mediated formation of a phosphodiester bond[7a] and the click dimerization of six-ring hairpin polyamides that bind to contiguous DNA sites.[7b] More recently, in the context of generating new DNA binders, Nagatsugi et al. have reported the use of dsDNA-templated click chemistry to generate acridine–peptide conjugates, which display slightly higher DNA affinity than the monomeric precursors.[7c]

Based on the well-known ability of distamycin-related polyamides to selectively bind to dsDNA sequences through antiparallel 2:1 insertion into the dsDNA minor groove,[8] and the recognition diversity that can be engineered by combining different aromatic amino acids like *N*-methylpyrrole (Py) or *N*-methylimidazole (Im),[9] we envisioned the possibility of using specific double-stranded oligonucleotides to control and program the coupling between designed heterocyclic polyamides.

We have previously shown that it is possible to accelerate a native chemical ligation (NCL) reaction between tripyrroles **2a**

* Corresponding author. E-mail address: joseluis.mascarenas@usc.es (J.L. Mascareñas).

0040-4020/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.tet.2013.05.126

and **3a**, which are, respectively, equipped with a thioester and a cysteine unit, by using a ds-oligonucleotide featuring a sequence that favors their homodimeric insertion in the minor groove (ATTTTA).[10] On the basis of these results we decided to study the effect of replacing one *N*-methylpyrrole unit of the polyamide by an *N*-methylimidazole. Hypothetically, this replacement should affect the DNA binding preferences of the resulting polyamide, and thus modify the sequence of the ds-oligonucleotides that best promote the reaction. This would represent a step forward toward encoding reactivity with specific dsDNAs.

Herein we demonstrate the viability of this concept by showing that the dsDNAs promoting the heterodimeric coupling are different than those inducing the formation of tripyrrole homodimers. Furthermore, we show that the reaction products of a mixture containing three different reactants depend on the sequence of the dsDNAs used as templates.

2. Results and discussion

The design of our reactive units relies on the well-known insertion mode of distamycin derivatives in the minor groove of dsDNA, which places the methyl groups of the pyrroles facing the solvent exposed surface of the complex.[11] Therefore, considering a 2:1 binding mode, these positions provide suitable linking points to introduce the reactive groups required for the coupling reaction (Fig. 1). The increased effective molarity provided by the template should accelerate the coupling, provided that the tether exhibits a suitable length and flexibility. As coupling reaction we chose a native chemical ligation (NCL) as it is compatible with aqueous media, and does not require additives or special pH or salt conditions.[12]

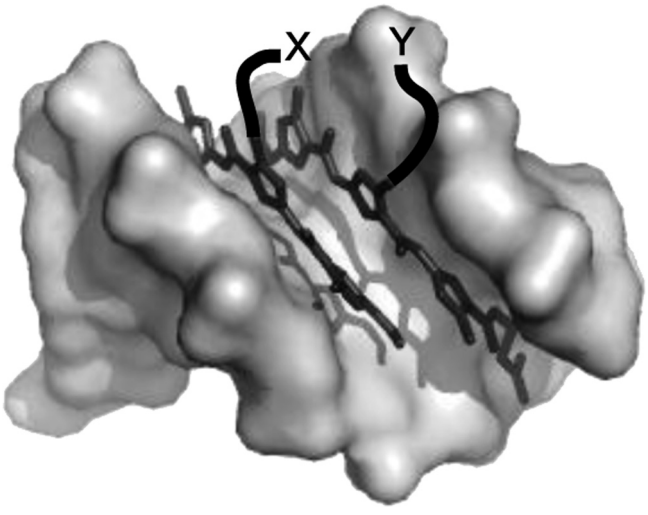


Fig. 1. Model of the location of reactive groups X and Y introduced in the central pyrrole of polyamides interacting with the DNA minor groove (built on a reported X-ray structure: Pdb 378d).[11]

The synthesis of the polyamide fragments was readily achieved following the protocols indicated in Scheme 1. The required tripyrrolic (PyPyPy) precursors **1a** and **1b** were assembled in four and three steps (34% and 38% overall yields), respectively, from commercially or readily available pyrrole and imidazole units (see Supplementary data). Alkylation of these precursors with *p*-tolyl 2-bromoethanethioate gave the thioesters **2a** and **2b**, which would be used as electrophilic partners in the NCL reaction. In our previous studies we had prepared a similar phenyl thioester that lacks the methyl substitution at the phenyl ring; we have now found that the introduction of this methyl substitution is convenient because it increases the stability of the thioester toward hydrolysis, and facilitates the subsequent coupling studies. The tripyrrole counterpart, equipped with the nucleophilic cysteine (**3a**) was prepared as shown in Scheme 1, by an initial alkylation of **1a** with *tert*-butyl (3-iodopropyl)carbamate, followed by deprotection, coupling of the amine with a trityl cysteine amino acid, and removal of the trityl group. If instead of coupling the cysteine to the intermediate amine, this is treated with 5-oxo-5-(phenylthio)pentanoic acid, we obtain the thioester **4a**, which represents an interesting probe to test the effect of the tether length in the coupling reaction.

In the case of the dipyrrole-imidazole system (ImPyPy), the cysteine-equipped partner had to be synthesized using an alternative strategy (bottom part of Scheme 1), because in this case the coupling of the iodo-propylamine to **1b** was low yielding. The compound was efficiently prepared by first introducing the side chain in the dipyrrole, to give **5** (see Supplementary data), followed by coupling of the imidazole, and introduction of the cysteine unit in the pyrrole tether (Scheme 1, bottom).

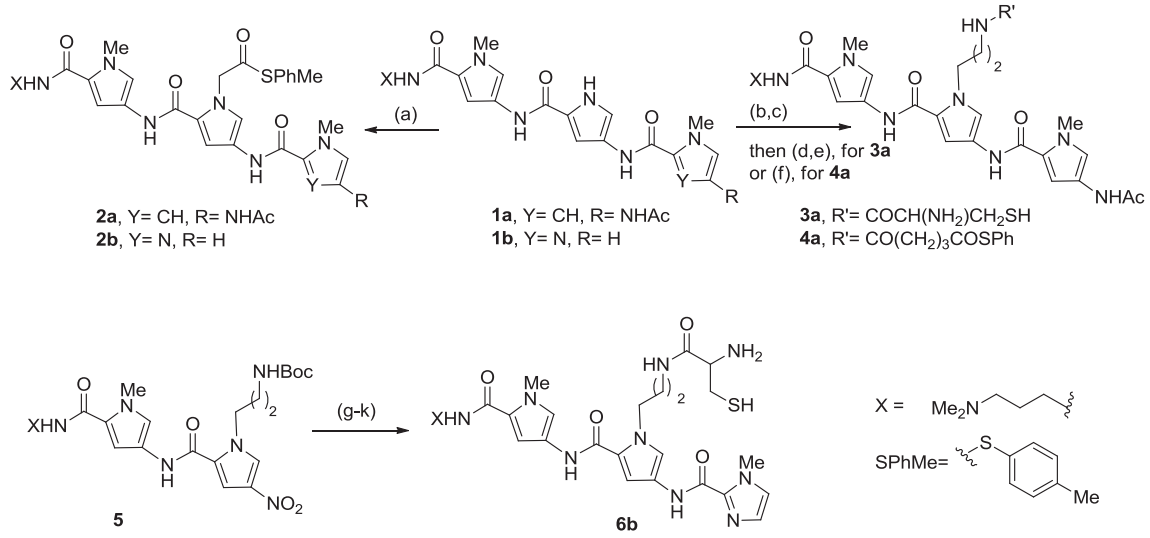
The NCL proceeds through an initial trans-thioesterification between the electrophilic partner (thioester) and the nucleophile (cysteine derivative) to give an intermediate (**I**), which evolves to the final amide by means of an intramolecular acyl transfer (Scheme 2).[12]

In agreement with previous results,[10] mixing the tripyrroles **2a** (5.0 μM) and **3a** (5.0 μM) in a 200 mM NaCl, 20 mM phosphate buffer containing 5 mM TCEP, at pH 7.0 and ~22 °C, does not generate significant amounts of the coupling product after 20 min. However, in the presence of 1 equiv of a hairpin ds-oligonucleotide (**A**) featuring an ATTTTA sequence, there is a significant acceleration in the reaction, as deduced from the HPLC traces of Fig. 2. Considering both the intermediate **Iaa** and the final product **7aa**, we have calculated an approximate conversion of 85–90% after 8 h.

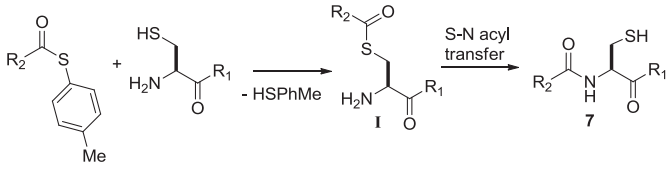
To study the effect of different DNAs, we used more diluted solutions (0.65 μM of the reactants, [phosphate]=12 mM, [NaCl]=120 mM, [TCEP]=3 mM, pH 7.0, and ~22 °C); and the results are represented in Fig. 2b. In terms of initial rates, the reaction in the presence of the DNA **A** is accelerated over 40 times with respect to that in its absence. An oligonucleotide featuring a guanine in the middle of the A/T tract (DNA **B**) is also capable of accelerating the reaction, which is in consonance with the known ability of this type of sequences to accept the antiparallel insertion of two polyamides in its minor groove.[8b] Oligonucleotides that are known to favor 1:1 over 2:1 insertions, like DNA **C** are significantly poorer catalysts. The reaction is also accelerated with sub-stoichiometric amounts of the oligonucleotides while using 0.5 equiv of the oligo **A** (see Supplementary data) generates almost similar results than with 1 equiv, further decrease in the amount of the catalyst leads to slower reaction rate and yield.[10]

It was also interesting to observe that the thioester **4a**, featuring a longer linking tether, leads to inferior results (curve x, Fig. 2b), suggesting a major entropic cost for the coupling reaction as a consequence of the tether length.

With these data at hand, we explored the heterodimeric couplings between the tripyrroles (PyPyPy) and the polyamides featuring an N-terminal *N*-methylimidazole instead of an *N*-methylpyrrole (ImPyPy). As shown in the left part of Fig. 3, the reaction between the tripyrrole-cysteine derivative **3a** and the imidazole-dipyrrole thioester **2b** gives the corresponding intermediate **Iab** and reaction product **7ab**, when carried out in the presence of 1 equiv of the hairpin DNA **B**. The effects of the different DNAs in the reaction rate are represented in Fig. 3b. It is important to note that in this case, and in contrast to the coupling of two tripyrroles, the reaction is accelerated only by the oligo **B** (ATGTTA, over 23-fold rate enhancement) and not by the oligo **A** (ATTTTA), which seems consistent with the known preferences of insertion of



Scheme 1. Synthesis of the reactive partners. Reactions: (a) $BrCH_2COSPhMe$, K_2CO_3, acetone/DMF, rt, 55% (for **2a**), 35% (for **2b**); (b) **1a**, $I(CH_2)_3NHBoc$, K_2CO_3, acetone/DMF (6:1), reflux, 38%; (c) TFA/CH_2Cl_2, 0 °C to rt; (d) Boc–Cys(Trt)–OH, HATU, DIPEA, CH_2Cl_2, rt; (e) TFA/Et_3SiH, CH_2Cl_2, 0 °C to rt, 48% (three steps); (f) $HOOC(CH_2)_3COSPh$, HOBt, HBTU, DIPEA/DMF, rt, 49%; (g) H_2, Pd/C 10%; (h) 1-methyl-1*H*-imidazole-2-carboxilic acid, DECP, Et_3N, DMAP, −10 °C, 44% (two steps); (i) TFA/CH_2Cl_2, 0 °C to rt; (j) Boc–Cys(Trt)–OH, HATU, DIPEA, CH_2Cl_2, rt; (k) TFA/Et_3SiH, CH_2Cl_2, 0 °C to rt, 43% (three steps).



Scheme 2. The native chemical ligation (NCL) involves the initial formation of thioester intermediate **I**.

Considering the above results, which confirmed that introducing the N-terminal imidazole unit in the pyrrole-based polyamides changes the sequence of the dsDNAs templates that best promote the coupling process; we wondered whether given a mixture of different monomeric reactants, we could alter the reaction profile by using different dsDNAs. In consonance with the reaction rates previously observed, if we add 1 equiv of the hairpin **A** (ATTTA) to an equimolecular mixture of tripyrrole **3a** and thioesters **2a** and **2b**, we observe almost exclusively the formation of the homodimeric product (**Iaa** plus **7aa**, brown peaks in Scheme 3), and only traces of the intermediate **Iab** (green peak). On the other hand, if we add the DNA **B** (ATGTTA), in addition to the homodimer we also observe the formation of significant amounts of the heterodimer (**Iab** and **7ab**, green peaks). Although, of course, still far from the objective of getting a complete change in selectivity in function of the DNA inputs, these results demonstrate the viability of the concept, and set the basis for future work toward this goal.

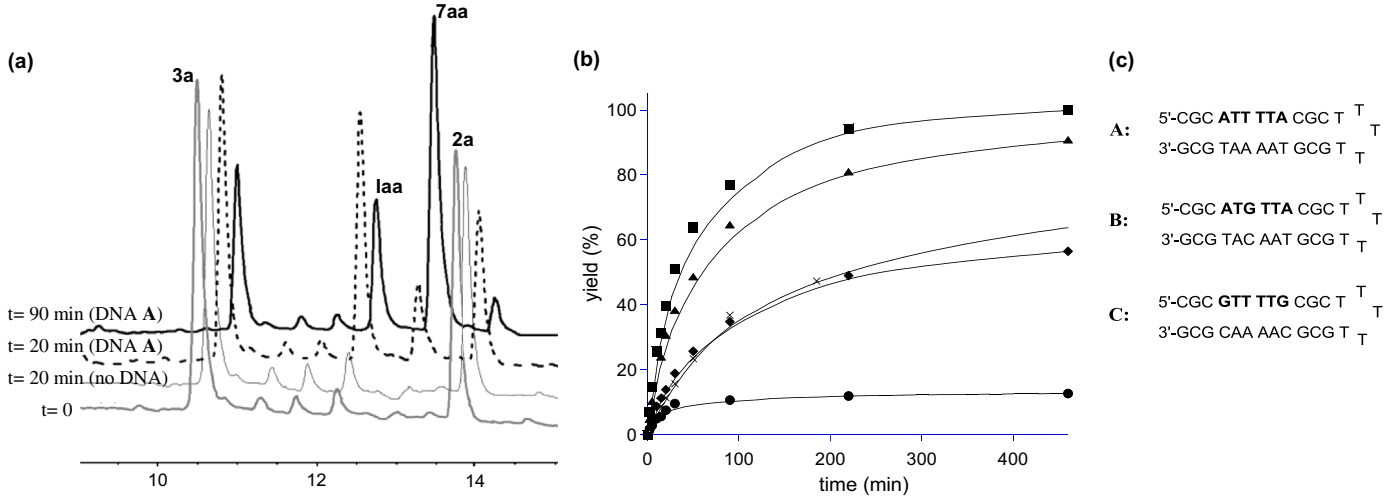


Fig. 2. (a) HPLC trace of the reaction between **2a** and **3a**, in the absence and in the presence of oligonucleotide **A**. (b) Graphic showing the formation of NCL products (**Iaa**+**7aa**) as function of time in the reaction of **2a** and **3a** (0.65 μM each, 12 mM phosphate, 120 mM NaCl, pH 7.0, 3 mM TCEP, ~22 °C; average of three experiments) in the absence of DNA (●), and in the presence of DNAs **A** (▲), **B** (■) and **C** (◆). We also show the reaction of **3a** and **4a** in the presence of DNA **A** (×). (c) Hairpin oligonucleotides used for the study.

imidazole-containing polyamides in minor grooves featuring a guanine step. We have also found that the alternative heterodimeric coupling, using the $Im(Py)_2$ unit as nucleophile (**6b**) and the tripyrrole as electrophile (**2a**), exhibits a similar acceleration profile in the presence of the different DNAs (see Supplementary data). Therefore, although in the case of the heterodimer coupling the rate increase induced by the DNA is less pronounced than for the formation of homodimers, the reaction is only accelerated by the DNA **B**.

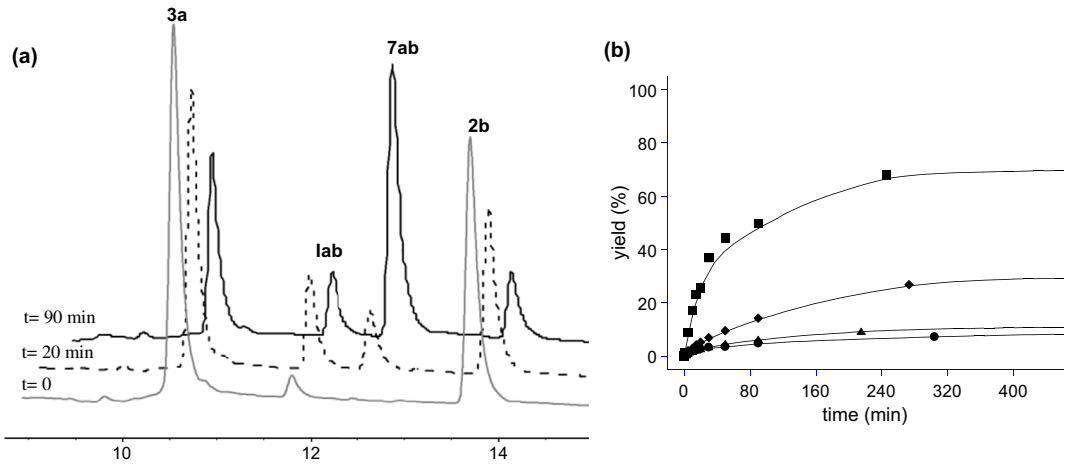
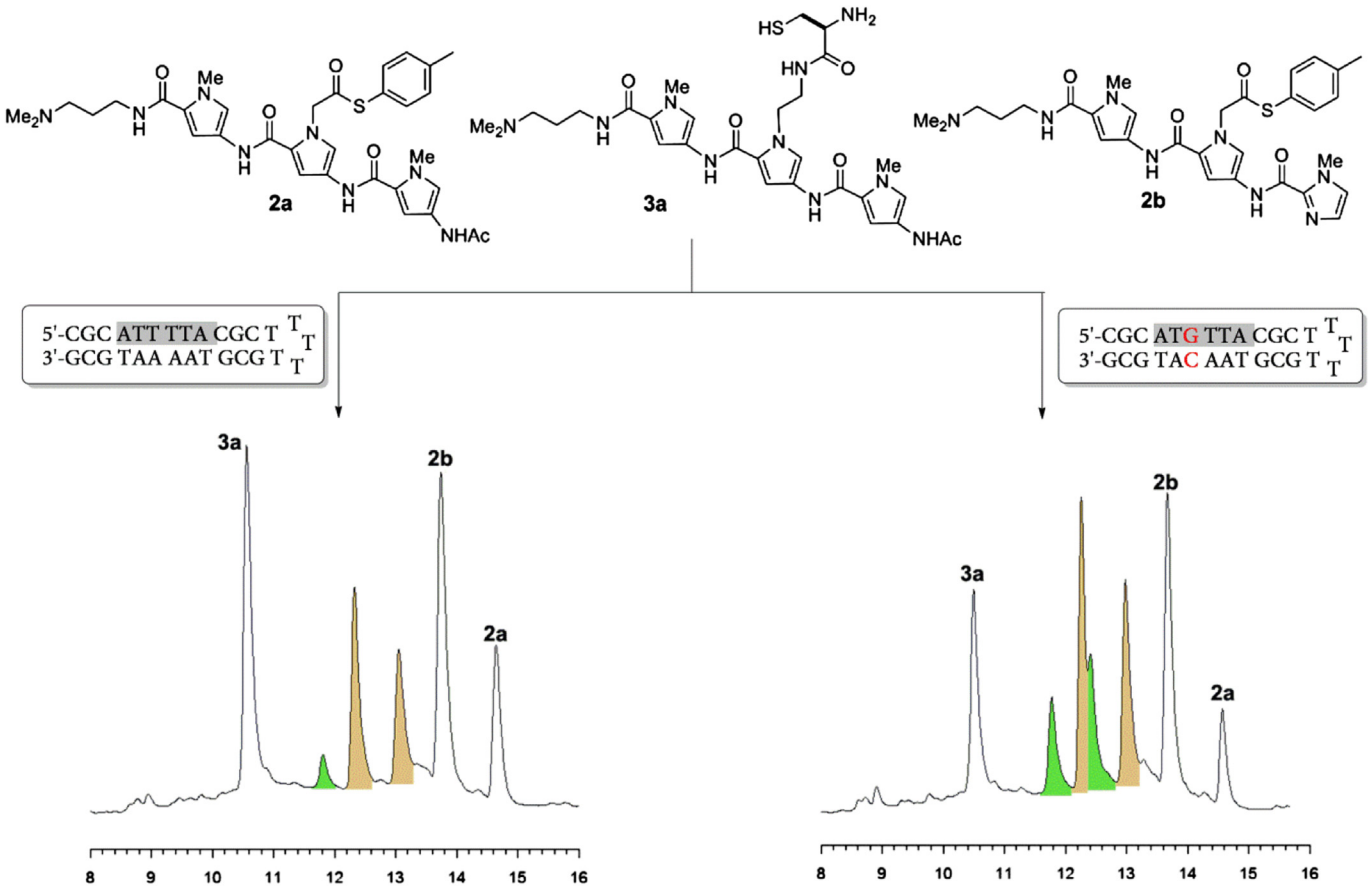


Fig. 3. (a) HPLC trace of the reaction between cysteine derivative **3a** and thioester **2b** in the presence of oligo **B** (conditions: 5.0 μM of each reactant, 200 mM NaCl, 20 mM phosphate buffer at pH 7.0, 5 mM TCEP, ~22 °C). (b) Graphic showing the formation of NCL products (**Iab**+**7ab**) as function of time (0.65 μM each reactant, 12 mM phosphate, 120 mM NaCl, pH 7.0, 3 mM TCEP, ~22 °C) in the absence of DNA (●), and in the presence of DNAs **A** (▲), B (■), and C (◆).



Scheme 3. Different outcome of the reaction of **3a** with the thioester partners **2a** and **2b**, depending on the DNA promoter. Conditions: 0.65 μM of reactants, 12 mM phosphate buffer, 120 mM NaCl, pH 7.0, 3 mM TCEP, 30 min of reaction, ~22 °C. Marked in brown the intermediate **Iaa** and product **7aa**, and in green the corresponding products **Iab** and **7ab**.

3. Conclusion

In summary, we have demonstrated that the rate of an NCL between appropriately functionalized DNA binding heterocyclic polyamides can be considerably increased by using specific double-stranded DNAs. While a dsDNA featuring an ATTTTA site promotes homodimeric coupling of tripyrroles (PyPyPy), but not the formation of heterodimers between tripyrroles and a polyamide with two pyrroles and one imidazole unit (ImPyPy), a dsDNA with an ATGTTA sequence accelerates both coupling reactions. Therefore, given a mixture of the three coupling partners, the reaction outcome is different depending on the sequence of the template DNA. Although the strategy is yet to be optimized, our results suggest the viability of using dsDNAs to encode reactivity. Potential advantages of using dsDNA over ssDNA as reactivity templates stem from the possibility of inducing asymmetric transformations because of the intrinsic chirality of the dsDNA scaffold, the potential applications for encoding synthetic processes in natural settings without the

requirement of separating the DNA strands, and the presumably easier implementation of catalytic chemistry.

4. Experimental section

4.1. General

Dimethylformamide (DMF) and trifluoroacetic acid (TFA) were from Scharlau, CH_3CN from Merck, CH_2Cl_2 from Panreac, the Boc–Cys(Trt)–OH from NovaBioChem, and the coupling agents HOBt (1-hydroxybenzotriazole), HBTU (*O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyluroniumhexafluorophosphate) and HATU (2-(1*H*-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluroniumhexafluorophosphate methanaminium) from GL Biochem (Shanghai). All other chemicals were purchased from Sigma–Aldrich or Fluka. All solvents were dry and synthesis grade, unless specifically noted. Oligonucleotides were acquired from Thermo-Fischer Sci.

Reactions were performed under Ar, unless specifically stated. External bath thermometers were used to record all reaction temperatures. Thin-layer chromatography (TLC) was performed on F254 250-μm silica gel plates or 79–230 mesh aluminum oxide, and components were visualized by observation under UV light, or by treating the plates with ninhydrine or cerium nitrate followed by heating.

Some reactions were followed by analytical reverse-phase HPLC (High-Performance Liquid Chromatography) with an Agilent 1100 series LC/MSD (G1956A Quad VL model; ESI-quadrupole, positive scan) using a Zorbax Eclipse XDB-C8 (5 μm, 4.6×150 mm) analytical column from Agilent. Standard conditions for analytical RP-HPLC consisted on a linear gradient from 5% to 95% of solvent B for 30 min at a flow rate of 1 mL/min (solvent A: water with 0.1% TFA, solvent B: acetonitrile with 0.1% TFA). Compounds were detected by UV at 220 and 304 nm.

NMR spectra were registered with Brucker WM-250, or with Varian Inova 400 and Inova 500 spectrometers, at rt.

MS spectra were acquired using a Micromass Autospec (for CI and FAB), an HP-5972-MSD (for EI), and both Bruker Biotof II and micrOTOF (for ESI).

When necessary, HPLC purifications were performed in an Agilent 1100 diode array HPLC system and using as column packing Nucleosil 120-10 C-18 (250×8 mm) for semipreparative purifications. Mixtures of CH_3CN (B) and H_2O (A) with 0.1% TFA were used as mobile phase with linear gradients. Detection was made at 220 and 304 nm simultaneously. Water was purified using a Millipore MilliQ water purification system.

4.2. Synthesis

*4.2.1. N2-[3-(Dimethylamine)propyl]-1-methyl-4-[(1-[2-oxo-2-(p-tolylthio)ethyl]-4-{[1-methyl-4-(methylcarboxamide)-1H-2-pyrrolyl] carboxamide}-1H-2-pyrrolyl)carboxamide]-1H-2-pyrrolcarboxamide (**2a**).* Tripyrrole **1a** (38.0 mg, 0.077 mmol) was added to a suspension of dry K_2CO_3 (159 mg, 1.150 mmol) in DMF (0.65 mL). A solution of *S*-*p*-tolyl 2-bromoethanethioate (112.6 mg, 0.460 mmol) in 2 mL of dry acetone was added and the mixture stirred for 30 min at rt. The solid was removed by filtration through Celite, the filtrate was concentrated, and the resulting residue was purified by RP-HPLC (H_2O/CH_3CN) to give **2a** as an amorphous yellow-brown solid (28 mg, 55%).

^{1}H NMR (250 MHz, CD_3OD) δ (ppm): 7.26–6.80 (m, 10H); 4.63 (s, 6H); 4.41 (s, 2H); 3.80 (s, 3H); 3.60 (s, 2H); 3.31 (s, 2H); 3.10 (s, 6H); 2.27 (s, 3H); 2.02 (s, 3H); 0.93 (m, 2H). ^{13}C NMR (62.9 MHz, CD_3OD) δ (ppm): 192.5 (C); 171.3 (C); 164.2 (C); 161.5 (C); 160.8 (C); 136.1 (C); 132.4 (CH); 131.9 (C); 131.0 (CH); 125.3 (C); 124.7 (C); 124.6 (C); 123.3 (C); 123.2 (C); 120.5 (CH); 114.7 (CH); 106.1 (CH); 105.9 (CH); 104.6 (CH); 68.7 (CH_2); 65.2 (CH_2); 53.3 (CH_3); 38.5 (CH_3); 25.2 (CH_2); 22.9 (CH_3); 20.5 (CH_3). MS (ESI): $[MH]^+$ calcd for $C_{33}H_{41}N_8O_5S$=661.2, found: *m*/*z* 661.2. HRMS ESI Q-TOF: $[MH]^+$ calcd for $C_{33}H_{41}N_8O_5S$=661.2915, found: *m*/*z* 661.2919.

*4.2.2. N2-[3-(Dimethylamine)propyl]-1-methyl-4-[(1-[2-oxo-2-(p-tolylthio)ethyl]-4-{1-methyl-1H-imidazole-2-carboxamide}-1H-2-pyrrolyl)carboxamide]-1H-2-pyrrolcarboxamide (**2b**).* Polyamide **1b** (14.9 mg, 0.034 mmol) was added to a suspension of dry K_2CO_3 (70.1 mg, 0.508 mmol) in DMF (0.25 mL). A solution of *S*-*p*-tolyl 2-bromoethanethioate (49.8 mg, 0.203 mmol) in 0.8 mL of dry acetone was slowly added to the resulting solution and the mixture was stirred for 15 min at rt. The solid was removed by filtration through Celite, the filtrate was concentrated, and the resulting residue was purified by RP-HPLC (H_2O/CH_3CN) to give **2b** as an amorphous brown solid (7.2 mg, 35%).

^{1}H NMR (250 MHz, CD_3OD) δ (ppm): 7.26–6.75 (m, 10H); 4.82 (s, 6H); 4.65 (s, 2H); 3.98 (s, 3H); 3.86 (s, 3H); 3.36 (m, 2H); 3.21 (s, 2H); 3.15 (s, 3H); 1.98 (s, 2H). ^{13}C NMR (62.9 MHz, CD_3OD) δ (ppm): 170.4 (C); 164.3 (C); 163.5 (C); 162.3 (C); 160.8 (C); 143.1 (C); 136.4 (CH); 132.0 (CH); 126.3 (C); 125.7 (C); 125.6 (C); 125.2 (CH); 123.3 (C); 123.2 (C); 121.5 (CH); 116.7 (CH); 106.1 (CH); 105.9 (CH); 104.6 (CH); 66.7 (CH_2); 64.2 (CH_2); 53.3 (CH_3); 37.8 (CH_3); 37.5 (CH_3); 36.8 (CH_2); 25.2 (CH_2); 23.5 (CH_3). MS (ESI): $[MH]^+$ calcd for $C_{30}H_{37}N_8O_4S$=605.2, found: *m*/*z* 605.2. HRMS ESI Q-TOF: $[MH]^+$ calcd for $C_{30}H_{37}N_8O_4S$=605.2653, found: *m*/*z* 605.2642.

*4.2.3. N2-[3-(Dimethylamine)propyl]-1-methyl-4-[(1-{3-(2-amine-3-mercaptopropanamide) propyl}-4-{[1-methyl-4-(methylcarboxamide)-1H-2-pyrrolyl]carboxamide}-1H-2-pyrrolyl)carboxamide]-1H-2-pyrrolcarboxamide (**3a**).* Tripyrrole **1a** (50.4 mg, 0.102 mmol) was added to a suspension of dry K_2CO_3 (210 mg, 1.52 mmol) in 3 mL of dry acetone and DMF (0.5 mL). *tert*-Butyl (3-iodopropyl) carbamate (173.8 mg, 0.610 mmol) was added, and the resulting mixture was refluxed for 8 h. The solid was removed by filtration through Celite, the filtrate was concentrated, and the resulting residue was purified by RP-HPLC (H_2O/CH_3CN) to give an amorphous pale-brown solid (25 mg, 38%).

^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 7.31–6.85 (m, 6H); 4.29 (m, 2H); 3.91 (s, 3H); 3.69 (s, 3H); 3.35 (m, 2H); 2.78 (m, 2H); 2.54 (m, 2H); 2.40 (m, 2H); 2.34 (s, 6H); 1.80 (s, 3H); 1.42 (m, 2H); 1.04 (s, 9H). ^{13}C NMR (100.6 MHz, CD_3OD) δ (ppm): 173.3 (C); 167.2 (C); 164.4 (C); 163.7 (C); 150.8 (CH); 136.1 (C); 135.6 (C); 126.2 (C); 126.1(C); 120.4 (CH); 109.6 (CH); 109.1 (CH); 107.9 (CH); 107.6 (CH); 79.8 (C); 61.2 (CH_2); 48.2 (CH_2); 47.6 (CH_3); 41.0 (CH_3); 39.8 (CH_2); 39.7 (CH_3); 37.2 (CH_3); 31.6 (CH_2); 31.1 (CH_3); 28.2 (CH_3); 25.8 (CH_2). MS (ESI): $[MH]^+$ calcd for $C_{32}H_{48}N_9O_6$=654.7, found: *m*/*z* 654.3.

The resulting product (25 mg, 0.043 mmol) was dissolved in 1 mL of 20% TFA/CH_2Cl_2 at 0 °C and then stirred for 75 min and allowed to reach rt. The solvents were removed under reduced pressure and the resulting crude was re-dissolved in DIPEA (10.6 μL, 0.057 mmol) and 1 mL CH_2Cl_2 and cooled to 0 °C. A solution of HATU (18.2 mg, 0.048 mmol), DIPEA (8.8 μL, 0.051 mmol), and Boc–Cys(Trt)–OH (22.3 mg, 0.048 mmol) in 3 mL of CH_2Cl_2 at 0 °C was then added. The mixture was stirred for 30 min at 0 °C and then purified by RP-HPLC (H_2O/CH_3CN). Fractions containing the desired product were lyophilized and re-dissolved in a solution of Et_3SiH (1 μL) in 1 mL of 20% TFA/CH_2Cl_2 at 0 °C. The mixture was stirred for 45 min at rt, the solvents were removed under reduced pressure, and the resulting residue was purified by RP-HPLC (H_2O/CH_3CN) to give **3a** as an amorphous brown solid (13.4 mg, 48%).

^{1}H NMR (500 MHz, CD_3OD) δ (ppm): 7.00–6.82 (m, 6H); 3.89 (m, 2H); 3.69 (s, 3H); 3.67 (s, 3H); 3.31 (br s, 5H); 2.92 (s, 6H); 1.94 (s, 3H); 1.82–1.75 (br s, 4H); 1.15–1.07 (m, 4H). ^{13}C NMR (125.8 MHz, CD_3OD) δ (ppm): 171.3 (C); 169.2 (C); 164.7 (C); 164.5 (C); 161.0 (C); 130.9 (CH); 129.7 (CH); 127.7 (CH); 125.5 (CH); 125.0

(C); 124.4 (C); 123.3 (C); 122.9 (C); 63.9 (CH_2); 63.6 (CH_2); 56.6 (CH_3); 52.1 (CH); 37.9 (CH_2); 37.5 (CH_2); 37.3 (CH_3); 26.7 (CH_2); 24.7 (CH_2); 24.3 (CH_2); 22.9 (CH_3). MS (ESI): $[MH]^+$ calcd for $C_{30}H_{45}N_{10}O_5S$=657.3, found: *m*/*z* 657.3. HRMS ESI Q-TOF: $[MH]^+$ calcd for $C_{30}H_{45}N_{10}O_5S$=657.3290, found: *m*/*z* 657.3283.

*4.2.4. N2-[3-(Dimethylamine)propyl]-1-methyl-4-[(1-{3-(5-oxo-5-[phenylthio]pentanamide)propyl}-4-{[1-methyl-4-(methylcarboxamide)-1H-2-pyrrolyl]carboxamide}-1H-2-pyrrolyl)carboxamide]-1H-2-pyrrolcarboxamide (**4a**).* The crude amine (21 mg, 0.034 mmol) resulting after removal of the Boc protecting group, as described in the previous experiment, was re-dissolved in DMF (0.7 mL) and 0.5 M DIPEA/DMF (68 μL). The resulting solution was added to a previously prepared solution of 5-oxo-5-((phenylthio) oxy)pentanoic acid (11.33 mg, 0.0505 mmol) in 285 μL of 0.5 M DIPEA/DMF containing HBTU (19.1 mg, 0.0505 mmol) and HOBt (9.1 mg, 0.0674 mmol), and the mixture was stirred for 3 h at rt. The crude residue after concentration was purified by RP-HPLC (H_2O/CH_3CN) to give **4a** as an amorphous brown solid (12.45 mg, 49%).

^{1}H NMR (500 MHz, CD_3OD) δ (ppm): 7.33–6.75 (m, 11H); 4.48 (br s, 4H); 3.79 (m, 2H); 3.39 (m, 2H); 3.16 (s, 3H); 2.96 (s, 6H); 2.61 (m, 2H); 2.17 (m, 2H); 2.11 (s, 3H); 1.97 (s, 3H); 1.93 (m, 2H); 1.87–1.83 (m, 2H); 1.10–1.06 (m, 2H). ^{13}C NMR (125.8 MHz, CD_3OD) δ (ppm): 135.7 (CH); 130.6 (CH); 130.3 (CH); 106.0 (CH); 63.3 (CH_2); 51.6 (CH_3); 43.5 (CH_2); 37.0 (CH_2); 36.9 (CH_3); 35.7 (CH_2); 30.7 (CH_2); 30.0 (CH_3); 22.5 (CH_3); 22.4 (CH_2); 22.1 (CH_2). MS (ESI): $[MH]^+$ calcd for $C_{38}H_{50}N_9O_6S$=760.3, found: *m*/*z* 760.3. HRMS ESI Q-TOF: $[MH]^+$ calcd for $C_{38}H_{50}N_9O_6S$=760.3599, found: *m*/*z* 760.3593.

*4.2.5. N2-[3-(Dimethylamine)propyl]-1-methyl-4-[(1-{3-(2-amine-3-mercaptopropanamide) propyl}-4-{1-methyl-1H-imidazole-2-carboxamide}-1H-2-pyrrolyl)carboxamide]-1H-2-pyrrolcarboxamide (**6b**).* A solution of **5** (50 mg, 0.096 mmol) in MeOH (2 mL) was hydrogenated for 1.5 h over 10% palladium on charcoal (35 mg) at rt. The catalyst was removed by filtration through Celite and the filtrate concentrated. The residue was immediately dissolved in DMF (1.2 mL) and cooled to 0 °C. It was then slowly added (20 min) to a solution of 1-methyl-1*H*-imidazole-2-carboxylic acid (14.0 mg, 0.110 mmol), DECP (0.07 mL, 0.482 mmol), Et_3N (0.02 mL, 0.144 mmol), and DMAP (catalytic) in THF (1.0 mL) at −10 °C. The mixture was stirred for 13 h at −10 °C. The solvents were removed under reduced pressure and the resulting residue was purified by RP-HPLC (H_2O/CH_3CN) to give the expected amide (25.4 mg, 44%).

^{1}H NMR (400 MHz, ketone) δ (ppm): 7.51–6.90 (m, 6H); 4.19 (m, 2H); 3.93 (s, 3H); 3.00 (s, 3H); 2.01 (d, 2H, *J*=6.7 Hz); 1.40 (s, 6H); 1.36 (m, 2H); 1.34 (m, 2H); 1.23 (t, 4H, *J*=6.7 Hz); 0.56 (s, 9H). ^{13}C NMR (100.6 MHz, ketone) δ (ppm): 210.5 (C); 164.4 (C); 131.5 (CH); 129.7 (C); 129.0 (C); 127.1 (C); 124.0 (C); 123.2 (C); 109.7 (CH); 109.5 (CH); 83.1 (C); 70.0 (CH_2); 67.6 (CH_2); 60.2 (CH_2); 51.6 (CH_2); 47.5 (CH_3); 37.3 (CH_2); 33.6 (CH_3); 30.2 (CH_2); 20.6 (CH_3); 5.7 (CH_3). MS (ESI): $[MH]^+$ calcd for $C_{29}H_{43}N_9O_5$=598.3, found: *m*/*z* 598.3.

A solution of this amide (14.4 mg, 0.024 mmol) in 1 mL of 20% TFA/CH_2Cl_2 at 0 °C was stirred for 1.5 h and that temperature and overnight at rt. The solvents were removed under reduced pressure and the resulting crude was re-dissolved in DIPEA (8.0 μL, 0.045 mmol) and 1 mL CH_2Cl_2 and cooled to 0 °C. This solution was added to a previously prepared solution of HATU (10.6 mg, 0.028 mmol), DIPEA (5.2 μL, 0.03 mmol), and Boc–Cys(Trt)–OH (13 mg, 0.028 mmol) in 3 mL of CH_2Cl_2 at 0 °C. The mixture was stirred for 30 min at 0 °C and then purified by RP-HPLC (H_2O/CH_3CN). Fractions containing the desired product were lyophilized and then dissolved in a solution of Et_3SiH (5 μL) in 1 mL of 20% TFA/CH_2Cl_2 at 0 °C. The resulting mixture was stirred for 30 min allowing to reach rt, the solvents were concentrated under reduced pressure, and the crude was purified by RP-HPLC (H_2O/CH_3CN) to give **6b** as an amorphous yellow-brown solid (6.3 mg, 44%).

^{1}H NMR (500 MHz, CD_3OD) δ (ppm): 7.49–6.89 (m, 6H); 4.48 (m, 2H); 4.13 (s, 3H); 4.05 (d, 1H, *J*=6.8 Hz); 3.92 (s, 3H); 3.46 (t, 2H, *J*=6.8 Hz); 3.36 (m, 2H); 3.23 (t, 4H, *J*=7.3 Hz); 2.95 (d, 6H, *J*=7.3 Hz); 2.06 (m, 2H); 1.40–1.30 (m, 2H). ^{13}C NMR (125.8 MHz, CD_3OD) δ (ppm): 168.4 (C); 164.9 (C); 161.1 (C); 140.0 (C); 127.5 (CH); 127.1 (CH); 124.3 (C); 124.0 (C); 123.3 (C); 123.1 (C); 122.8 (C); 121.1 (CH); 120.0 (CH); 106.6 (CH); 106.5 (CH); 56.6 (CH_2); 56.4 (CH); 49.1 (CH_3); 47.2 (CH_2); 43.5 (CH_3); 37.9 (CH_2); 37.0 (CH_3); 36.7 (CH_2); 36.2 (CH_3); 32.4 (CH_2); 26.5 (CH_2); 26.3 (CH_2). MS (ESI): $[MH]^+$ calcd for $C_{27}H_{41}N_{10}O_4S$=601.3, found: 601.3; $[M_2H]^+$ calcd for $C_{54}H_{82}N_{20}O_8S_2$=1202.6, found: *m*/*z* 1201.8. HRMS ESI Q-TOF: $[MH]^+$ calcd for $C_{27}H_{41}N_{10}O_4S$=601.3027, found: 601.3029.

4.3. DNA-templated reactions

The concentration of the reactants was calculated by UV measurements in water at 20 °C in a Jasco V-630 spectrophotometer coupled to an ETC-717 Peltier, using a standard UV Hellma micro cuvette 108.002-QS (10 mm light path). Extinction coefficients (ε) were obtained for each compound in water: ε_{304}=34.000 M^{-1} cm^{-1} for tripyrrole substrates, ε_{304}=26.000 M^{-1} cm^{-1} for imidazole-dipyrrole substrates, and ε_{304}=60.000 M^{-1} cm^{-1} for the hexapyrrolic products or ε_{304}=50.000 M^{-1} cm^{-1} for the heterodimeric products.

Reverse-phase HPLC analysis was performed using a Merck LiChrospher WP 300 RP-18 (5 μm, 250×4 mm) column. Binary gradients of solvents A (H_2O, 0.1% TFA) and B (CH_3CN, 0.1% TFA) were employed at a flow rate of 1.0 mL/min. The injection volume was 200–100 μL. The HPLC peaks were detected by monitoring the UV absorbance at λ=304 nm and their assignments were confirmed by coupled MS spectrometry. Graphics were represented using the program Kaleidagraph (v 3.5 by Synergy Software).

The following methodology was applied on DNA-templated NCL studies: All the reactions were repeated twice using conditions *a* to identify final product and intermediate, and three times using conditions *b*, with and without DNA (1 equiv), in both cases at rt (~22 °C) in phosphate buffer (pH 7.0) with NaCl. A typical experiment consists on the preparation of a solution of the cysteine containing compound (in 50 mM TCEP) and TCEP in the corresponding buffer/water mixture. Then this mixture was stirred for 5 min at rt with or without the corresponding ds-oligonucleotide and NCL was initiated by adding the thioester compound. Aliquots at different times were quenched on 0.1% TFA, frozen, and kept at −20 °C for analysis.

- Conditions *a*: Final concentration of the species [thioester]=[cysteine derivative]=[ADN]=5.0 μM, [TCEP]=5.0 mM, [phosphate buffer]=20 mM, [NaCl]=200 mM in 1000 μL, pH 7.0. 100 μL aliquots were directly injected on HPLC–MS.
- Conditions *b*: slower kinetics. Final concentration of the species [thioester]=[cysteine derivative]=[ADN]=0.65 μM, [TCEP]=3.0 mM, [phosphate buffer]=12 mM, [NaCl]=120 mM in 3000 μL, pH 7.0. 200 μL aliquots were taken at different times.
- Condition *b* was used to evaluate ds-oligonucleotide template processes, yields were calculated using Abs_{UV} (λ_{max}=304 nm) for coupling products (final product and intermediate) applying previously described extinction coefficients, and considering the sum of both peak areas, the thioester intermediate **I** and the coupled products **7**. The approximate rate of the NCL reaction was calculated from the slope for initial times.

The coupled intermediates (**I**) and products (**7**) were identified by LC–MS: Compound **7aa**: ESI $[MH]^+$ calcd for $C_{56}H_{77}N_{18}O_{10}S$=1193.6, found=1193.5. UV (H_2O) $\lambda_{max,\ nm}$ (ε): 304 (60,000 M^{-1} cm^{-1}). Compound **7aa′** (resulting by the coupling of **4a** and **3a**): ESI $[MH]^+$ calcd for $C_{62}H_{88}N_{19}O_{11}S$=1306.7, found=1307.0. UV (H_2O) $\lambda_{max,\ nm}$ (ε): 304

(60,000 $M^{-1}\,cm^{-1}$). HRMS ESI Q-TOF: $[MH]^+$ calcd for $C_{62}H_{88}N_{19}O_{11}S$=1306.6613, found=1306.6592. Compound **7ab**: ESI $[MH]^+$ calcd for $C_{53}H_{73}N_{18}O_9S$=1137.6, found=1136.6. UV (H_2O) $\lambda_{max,\ nm}$ (ε): 304 (50,000 $M^{-1}\,cm^{-1}$). HRMS ESI Q-TOF: $[MH]^+$ calcd for $C_{53}H_{73}N_{18}O_9S$=1137.5511, found=1137.5440. Compound **7ba** (resulting by the coupling of **6b** and **2a**): ESI $[MH]^+$ calcd for $C_{53}H_{73}N_{18}O_9S$=1137.6, found=1138.8. UV (H_2O) $\lambda_{max,\ nm}$ (ε): 304 (50,000 $M^{-1}\,cm^{-1}$). HRMS ESI Q-TOF: $[MH]^+$ calcd for $C_{53}H_{73}N_{18}O_9S$=1137.5511, found=1137.5522.

Acknowledgements

We thank the financial support provided by the Spanish grants SAF2010-20822-C02, CTQ2012-31341, CSD2007-00006, Consolider Ingenio 2010 and the Xunta de Galicia GRC2010/12. A.J-B. thanks the Spanish Ministerio de Educación for his PhD fellowship.

Supplementary data

We present information on the synthesis of the precursors **1a**, **1b**, and **5**. We also include the graphics for the reaction between **2a** and **6b** and for the coupling between **2a** and **3a** using 0.5 equiv of DNA **A**. Additional characterization data, such as MS spectra, are also included. Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.tet.2013.05.126.

References and notes

1. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, 5th ed.; Garland Science: 2007, Chapters 5–7.
2. (a) Orgel, L. E. *Acc. Chem. Res.* **1995**, *28*, 109–118; (b) Luther, A.; Brandsch, R.; von Kiedrowski, G. *Nature* **1998**, *396*, 245–248.
3. (a) Erben, A.; Grossmann, T. N.; Seitz, O. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 4993–4997; (b) Li, X.; Liu, D. R. *Angew. Chem., Int. Ed.* **2004**, *43*, 4848–4870; (c) He, Y.; Liu, D. *J. Am. Chem. Soc.* **2011**, *133*, 9972–9975.
4. (a) Chen, X.-H.; Roloff, A.; Seitz, O. *Angew. Chem., Int. Ed.* **2012**, *51*, 4479–4483; (b) Grossmann, T. N.; Roglin, L.; Seitz, O. *Angew. Chem., Int. Ed.* **2008**, *47*, 7119–7122; (c) Grossmann, T. N.; Seitz, O. *Chem.—Eur. J.* **2009**, *15*, 6723–6730.
5. Grossmann, T. N.; Strohbach, A.; Seitz, O. *ChemBioChem* **2008**, *9*, 2185–2192.
6. (a) Latchman, D. *Gene Control*; Garland Science, Taylor and Francis Group: New York, London, 2010; (b) Latchman, D. *Eukaryotic Transcription Factors*, 5th ed.; Elsevier/Academic Press: New York, 2008.
7. (a) Luebke, K. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 8733–8735; (b) Poulin-Kerstien, A. T.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 15811–15821; (c) Imoto, S.; Hirohama, T.; Nagatsugi, F. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5660–5663.
8. (a) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723–5727; (b) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235; (c) Vázquez, M. E.; Caamaño, A. C.; Mascareñas, J. L. *Chem. Soc. Rev.* **2003**, *32*, 338–349; (d) Pazos, E.; Mosquera, J.; Vázquez, M. E.; Mascareñas, J. L. *ChemBioChem* **2011**, *12*, 1958–1973.
9. (a) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284–299; (b) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468–471.
10. Dodero, V. I.; Mosquera, M.; Blanco, J. B.; Castedo, L.; Mascareñas, J. L. *Org. Lett.* **2006**, *8*, 4433–4436.
11. Mitra, S. N.; Wahl, M. C.; Sundaralingam, M. *Acta Crystallogr., Sect. D* **1999**, *55*, 602–609.
12. (a) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779; (b) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S. A.* **1999**, *96*, 10068–10073.