CHEMISTRY
A European Journal

A Journal of
ChemPubSoc Europe

Accepted Article

Title: Methylated nucleobases: Synthesis and evaluation for base pairing in vitro and in vivo

Authors: Amit M. Jabgunde, Faten Jaziri, Omprakash Bande, Matheus Froeyen, Mikhail Abramov, Hoai Nguyen, Guy Schepers, Eveline Lescrinier, Vitor B. Pinheiro, Valérie Pezo, Philippe Marlière, and Piet Herdewijn

To be cited as: *Chem. Eur. J.* 10.1002/chem.201802304

Link to VoR: http://dx.doi.org/10.1002/chem.201802304

Supported by
ACES

WILEY-VCH

10.1002/chem.201802304

WILEY-VCH

FULL PAPER

Methylated nucleobases : Synthesis and evaluation for base pairing *in vitro* and *in vivo*

Amit M. Jabgunde, [a] Faten Jaziri, [b] Omprakash Bande, [a] Matheus Froeyen, [a] Mikhail Abramov, [a] Hoai Nguyen, [a] Guy Schepers, [a] Eveline Lescrinier[a], Vitor B. Pinheiro,[c] Valérie Pezo,[b] Philippe Marlière,[b] and Piet Herdewijn*[a]

Abstract: The synthesis, base pairing properties and *in vitro* (polymerase) and *in vivo* (*E. coli*) recognition of 2'-deoxynucleotides with a 2-amino-6-methyl-8-oxo-7,8-dihydro-purine (**X**), a 2-methyl-6-thiopurine (**Y**) and a 6-methyl-4-pyrimidone (**Z**) base moiety are described. As demonstrated by T_m measurements, the **X** and **Y** bases fail to form a self-complementary base pair. Despite this failure, enzymatic incorporation experiments show that selected DNA polymerases recognize the **X** nucleotide and incorporate this modified nucleotide versus **X** in the template. *In vivo*, **X** is mainly recognized as a A/G or C base; **Y** is recognized as a G or C base and **Z** is mostly recognized as T or C. Replacing functional groups in nucleobases normally involved in W-C recognition (6-carbonyl and 2-amino group of purine; 6-carbonyl of pyrimidine) readily leads to orthogonality (absence of base pairing with natural bases).

Introduction

Expanding the genetic alphabet to include an additional base pair has led to the replication of such a base pair in a bacterial cell,[1] and in the use of such a base pair to code for unnatural amino acids.[2] Numerous research groups have studied base modified nucleic acids but few of them have been evaluated *in vivo*.[2-3] The candidate base pairs are comprised of reorganized donor–acceptor hydrogen-bonding groups,[4] increased number of hydrogen bonds,[5] size expanded bases,[3b, 6] hydrophobic bases,[7] and metal-ion coordinating bases.[8] For example, the work of Hirao,[9] and Romesberg,[1-2] demonstrates that hydrophobic interactions may contribute to stable base pairing. The work of Benner is an example how reorganization of donor-accepter interactions may lead to orthogonal base pairing *in vitro*.[4c] But many others, likewise, contributed to this research field,[10] and this research has been the subject of many reviews.[11]

Introduction of a methyl group in the 5-position of 2'-deoxyuridine (giving thymidine) and in the 5-position of 2'-deoxycytidine (giving 5-methyl-2'-deoxycytidine) increases duplex stability of double stranded nucleic acids due to hydrophobic interactions with the neighbouring bases.[12] Here we have replaced functionalities of purine and pyrimidine bases, normally involved in base pairing, with a methyl group to investigate the influence on duplex stability and to investigate the potential to involve them in new base pairing systems. A second small hydrophobic functionality that is involved in this study is a thiocarbonyl group (replacing a carbonyl group), for its potential to interact with the methyl group present in opposite bases in a duplex. By introducing such methyl and thiocarbonyl group, we will make the major (or minor groove) more hydrophobic which may not be favourable for duplex stabilization and could contribute to orthogonality. However, there are several precedents of the involvement of a hydrophobic thiocarbonyl group in base pair stabilization and some natural ^{t}RNA contain sulphur modified nucleobases such as 2-thio-U, 4-thio-U, 2-thio-C.[13] Rappaport has studied the 6-thioguanine (**G$_S$**): 5-methyl-2-pyrimidone (**T^H**) base pair (Figure 1a).[14] The laboratory of Switzer has investigated the 2-thio-isoguanine (**iG$_S$**):5-methyl-isocytosine (**iC^{Me}**) base pair[15] (Figure 1b) and the 2-thio-isoguanine(**iG$_S$**):5-methyl-4-pyrimidone (**P**)[16] base pair (Figure 1c). In all these cases the thiocarbonyl group was placed opposite an NH_2 group or an H-atom and the stability of the base pair (**G^S**:**T^H**, **iG^S**:**iC^{Me}** and **iG^S**:**P**) was found to be similar to the canonical A:T pair.

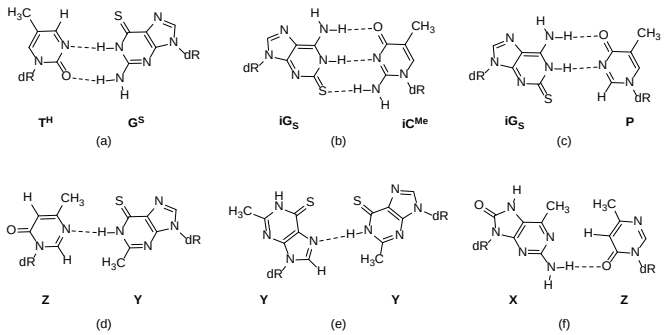


Figure 1. Schematic representation of putative base pairs between (a) 5-methyl-2-pyrimidone (**T^H**) with 6-thioguanine (**G^S**) (b) 2-thio-isoguanine (**iG$_S$**) with 5-methyl-isocytosine (**iC^{Me}**) (c) 2-thio-isoguanine (**iG$_S$**) with 5-methyl-4-pyrimidone (**P**) (d) 6-methyl-4-pyrimidone (**Z**) with 2-methyl-6-thiopurine (**Y**) (e) self-base pair of 2-methyl-6-thiopurine (**Y**:**Y**), (f) 2-amino-6-methyl-8-oxo-7,8-dihydro-purine (**X**) and 6-methyl-4-pyrimidone (**Z**)

[a] A. M. Jabgunde, O. Bande, M. Froeyen, M. Abramov, H. Nguyen, [G. Schepers, E. Lescrinier, V. Pezo, P. Marlière, P. Herdewijn
KU Leuven, Rega Institute, Medicinal Chemistry
Herestraat 49 box 1041, 3000 Leuven, Belgium
E-mail: Piet.Herdewijn@kuleuven.be
[b] Faten Jaziri, V. Pezo, P. Marlière
Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France
[c] V. B. Pinheiro,
Institute of Structural and Molecular Biology, University College London, Darwin Building, Gower Street, London
WC1E 6BT, United Kingdom

Supporting information for this article is given via a link at the end of the document.

10.1002/chem.201802304

The use of a self-complementary base for the expansion of the genetic alphabet has the advantage that the potential for mispairing with the natural bases is reduced (as only one base has to be considered for mispairing). It is not really a limitation since the addition of a single self-complementary base to the genetic alphabet would create 61 new codons. For example Pochet and Marlière[17] proposed a self-pair using a 2-amino-8-oxo-7,8-dihydro-purine base, with the potential to form two hydrogen bonds (Figure 2a). This work is based on the well-recognized ability of the mutagenic 8-oxoguanine to base pair with adenine base, the former oriented in the *syn* conformation. This 8-oxoG (*syn*):A(*anti*) base pair was studied by X-ray diffraction and NMR[18]

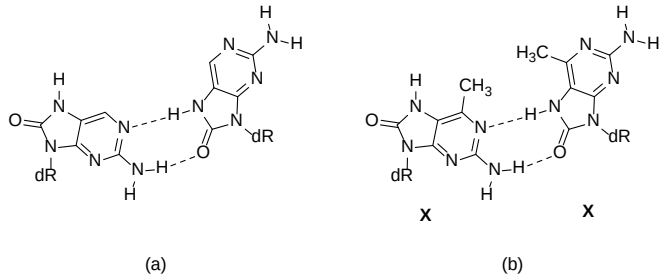


Figure 2. Schematic representation of a putative self–complementary base pair of (a) 2-amino-8-oxo-7,8-dihydro-purine (b) 2-amino-6-methyl-8-oxo-7,8-dihydro-purine (**X**).

Replacement of G with 8-oxoG increases thermal stability of duplex when A is placed in the opposite position.[19] The presence of the 8-oxo group induces glycosidic bond rotation which led to 8-oxoG(*syn*):A(*anti*) conformation.[18b] When 2-amino-8-oxo-7,8-dihydro-purine is placed opposite to A, thermal stability of the duplexes (compared with A:T) is decreased.[20] However, 2-amino-8-oxo-7,8-dihydro-purine:2-amino-8-oxo-7,8-dihydro-purine self-pair showed less destabilization. The relative stability of the 2-amino-8-oxo-7,8-dihydro-purine self-pair could be explained by the formation of two hydrogen bonds (Figure 2a).
A first example that we investigated is the potential self-complimentary base pair 2-amino-6-methyl-8-oxo-7,8-dihydro-purine (**X**) (Figure 2b) in the *syn*:*anti* conformation.

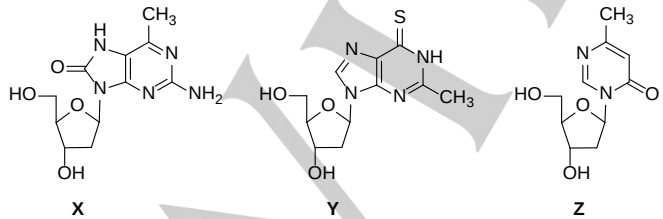


Figure 3. 2'-Deoxyribonucleosides, with modified purine (**X** and **Y**) and pyrimidine (**Z**) bases that have been used in this study. For the sake of simplicity, we use **X**, **Y** and **Z** as code, as well for the base as for the nucleos(t)ide.

Further, we have included in our study the 2-methyl-6-thiopurine base (**Y**) to pair with the 6-methyl-4-pyrimidone heterocycle (Figure 1d) (**Z**) and with itself (Figure 1e). In addition, potential pairing of the modified pyrimidone heterocycle (**Z**) with the previously mentioned 2-amino-6-methyl-8-oxo-7,8-dihydro-purine base (**X**) was investigated (Figure 1f).
There are many precedents in the literature of the stabilization of macromolecular structure or small molecule interactions by thiocarbonyl and methyl groups.[21] In addition the higher polarizability of sulphur over oxygen may contribute to enhanced stacking interactions.[22]
Our investigation started by building a static *in silico* model to analyse if **X:X** interactions and **Z:Y** interactions are potentially accessible in a dsDNA context. The models are shown in Figures 4 and 5. Although distances between functional groups are not optimal, closer contacts are possible in the pyrimidine:purine system (Figure 5) than in the purine self-pair system (Figure 4). The 6-methyl-4-pyrimidone nucleoside was modelled in the trans configuration, while the presence of the 4-carbonyl group would suggest that the cis configuration is more stable. The models demonstrate that, in theory, accommodation of both base pairs in a dsDNA structure might be possible.

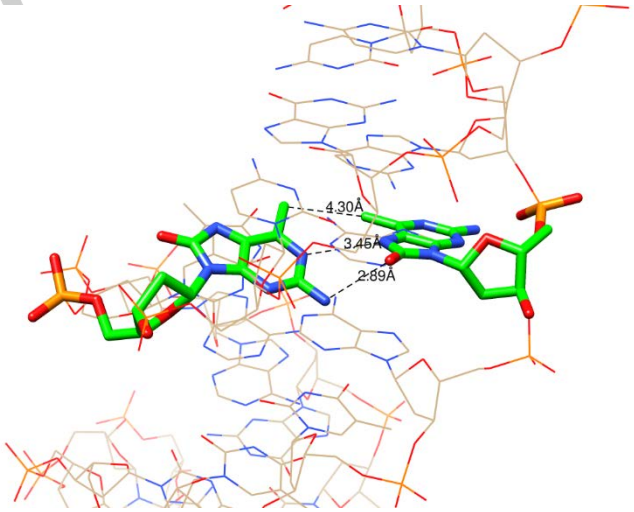


Figure 4. Static model showing potential base:base interactions in a 2-amino -6-methyl-8-oxo-7,8-dihydro-purine (**X**) self-pair.

10.1002/chem.201802304

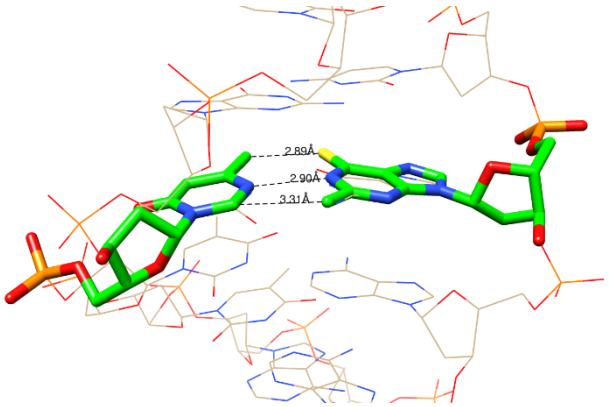
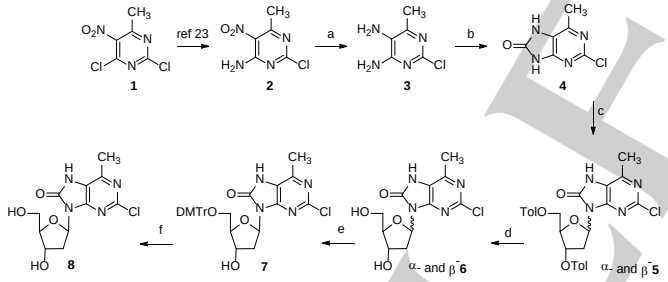


Figure 5. Static model showing potential base:base interactions in the 6-methyl-4-pyrimidone (**Z**) and 2-methyl-6-thiopurine (**Y**) base pair.

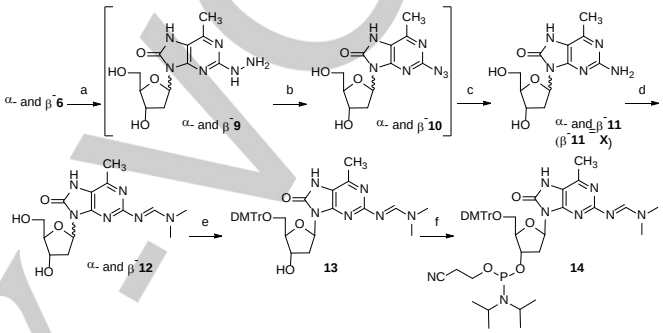
Result and discussion

The 2-chloro-4-amino-5-nitro-6-methyl pyrimidine **2** was synthesized according to a reported procedure with minor modifications.[23] The nitro group of **2** was reduced using sodium hydrosulphite to afford 2-chloro-4,5-diamino-6-methyl pyrimidine **3**. Reaction of 2-chloro-4,5-diamino-6-methyl- pyrimidine **3** with 1,1'-carbonyldiimidazole in 1,4-dioxane at 110 °C afforded compound **4**.



Scheme 1. Synthesis of 2-chloro-6-methyl-8-oxo-7,8-dihydro-purine nucleoside. (a) $Na_2S_2O_4$, H_2O:EtOAc:THF (1:1:3.5), rt, 3h, 39% (b) 1,1'-carbonyldiimidazole, 1,4-dioxane, 110 °C, 50 min, 86% (c) NaH, 1-chloro-2-deoxy-3,5-di-*O*-p-toluoyl-α-D-erythro-pentofuranose, 50 °C, 3h, 30%; (d) NaOMe, MeOH, rt 1h, 95%; (e) DMTrCl, pyridine, 0 °C to rt, 12 h, 58% β isomer; (f) 1.25 M HCl, MeOH:CH_2Cl_2 (1:1).

The sodium salt of **4**, produced in situ using NaH in acetonitrile was reacted with 1-chloro-2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-*erythro*-pentofuranose at 50 °C. The reaction mixture was purified on silica gel column to afford 2-chloro-6-methyl-8-oxo-7,8-dihydro-9-(2-deoxy-3,5-di-*O*-*p*-toluoyl-D-*erythro*-pentofuranos-1-yl)-purine (**5**) as an inseparable mixture of α and β-anomer in 30% yield. Treatment of **5** with sodium methoxide in methanol deprotected the toluoyl groups. The crude product was then purified by flash column chromatography to yield an inseparable anomeric mixture of **6** as a white solid. Compound **6** was converted to its dimethoxytrityl derivative **7** by selective protection of the primary alcohol with 4,4'-dimethoxytrityl chloride (DMTrCl) in pyridine. Flash column chromatography was used to separate the desired β-isomer from the α-isomer and ROESY and HMBC of 2D NMR was used to prove the correct structure. Trityl deprotection of compound **7** afforded the nucleoside **8**. 2D NMR analysis of compound **8** confirmed the previous structural analysis. However, amination of 2-chloro of compound **6** using ammonia under heating conditions was unsuccessful. Therefore, we decided to replace first the 2-chloro by a 2-azido group, and to separate the isomers after introduction of the 2-amino function.

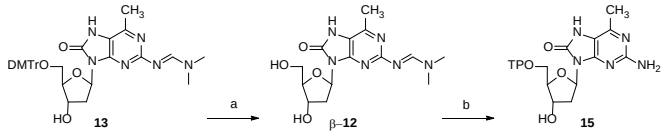


Scheme 2. Synthesis of the protected phosphoramidite of 2-amino-6-methyl-8-oxo-7,8-dihydro-9-(2-deoxy-β-D-*erythro*-pentofuranos-1-yl)-purine. (a) NH_2-NH_2, EtOH, 84 °C, 12 h; (b) $NaNO_2$, AcOH, –5 °C - rt, 12 h; (c) H_2, PtO_2, MeOH, 55 °C, 45 h, 21% over 3 steps; (d) DMF–DMA, MeOH, rt, 1 h; (e) DMTrCl, pyridine, 0 °C to rt, 12 h, 58% β anomer; (f) $(i\text{-}Pr_2N)_2POC_2H_4CN$, 1H-tetrazole, CH_2Cl_2, 0 °C to rt, 1 h, 64%.

Thus, nucleoside **6** was treated with hydrazine in ethanol to afford the 2-hydrazino derivative **9** which on further reaction with nitrous acid at –5 °C, led to formation of 2-azido-6-methyl-8-oxo-7,8-dihydro-purine nucleoside **10**. Reduction of azido group followed by N,N-dimethylaminomethylidene protection of the resulting amine function afforded compound **12** which, upon tritylation, gave a separable mixture of the α,β-anomers **13**. The desired β-anomer was isolated by flash column chromatography and converted into phosphoramidite **14** by reaction with bis(diisopropylamino)(2-cyanoethoxy)phosphine in the presence of tetrazole.

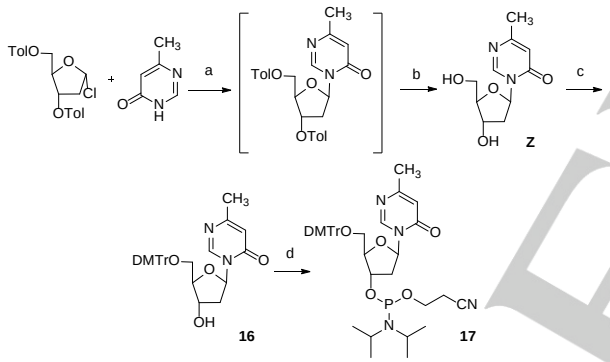
The triphosphate (of compound β-**12**) was synthesized according to the method described by Ludwig[24] starting from corresponding N,N-dimethylaminomethylidene protected nucleoside β-**12** (obtained by detritylation of compound **13**) (Scheme 3). In this procedure, regioselective phosphorylation of the 5'-hydroxy group of the sugar moiety was carried out with phosphoryl oxychloride in trimethyl phosphate followed by the addition of tetrabutylammonium pyrophosphate. The reaction product was deprotected with aqueous ammonia, isolated on Sephadex A25 column and final purification was performed by RP HPLC.

10.1002/chem.201802304



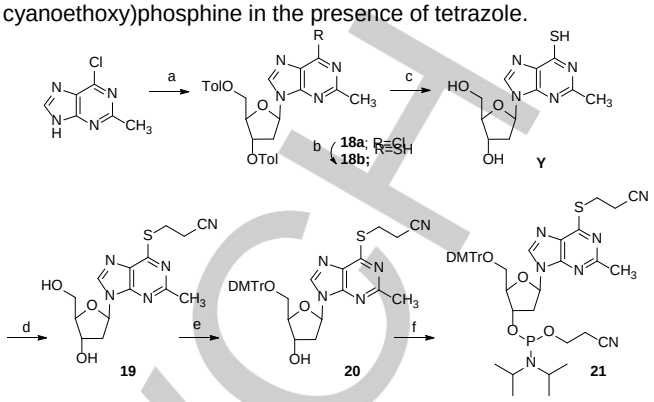
Scheme 3. Synthesis of the triphosphate **15** a) 1% TFA on Silica column b) i) trimethylphosphate, $POCl_3$, 0°C, 5 h; ii) Bu_3N, $(NBu_4)_3HP_2O_5$, 30 min; iii) 25% NH_3, 2 h;

As depicted in scheme 4, glycosylation of commercially available 6-methyl-4-pyrimidone with 3,5-bis-(toluoyl)-2-deoxyribosyl chloride using bis(trimethylsilyl)acetamide and $SnCl_4$ in dry acetonitrile afforded a crude mixture of 6-methyl-4-pyrimidone nucleoside. The cleavage of toluoyl-protecting group was carried out using 5 M solution of sodium methoxide in methanol to afford pure 6-methyl-3-(2-deoxy-β-D-*erythro*-pentofuranos-1-yl)-4-pyrimidone **Z** in 36% yield over two steps. Selective protection of the primary alcohol with 4,4'-dimethoxytrityl chloride (DMTrCl) in pyridine gave **16**, which was then converted to phosphoramidite **17** by reaction with bis(diisopropylamino)(2-cyanoethoxy) phosphine in the presence of tetrazole.



Scheme 4. Synthesis of the protected phosphoramidite of 6-methyl-3-(2-deoxy-5-*O*-DMTr-β-D-*erythro*-pentofuranos-1-yl)-4-pyrimidone. (a) $SnCl_4$, BSA, CH_3CN, rt, 2 h; (b) NaOMe, MeOH, 1.5 h, rt; (c) DMTrCl, pyridine, 0 °C to rt, 12 h, 69 %; (d) bis-(diisopropylamino)(2-cyanoethoxy)-phosphine, tetrazole, CH_2Cl_2, 0 °C to rt, 2.5 h, 62%.

The sodium salt of 2-methyl-6-chloropurine, produced in situ using NaH in acetonitrile, was reacted with l-chloro-2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-*erythro*-pentofuranose at room temperature to afford the desired nucleoside **18a** in 76% yield (Scheme 5). The 6-chloro group in **18a** was converted into a thiol group by treatment with sodium thiosulfate. Cleavage of toluoyl protecting groups by a 5 M solution of sodium methoxide in methanol (compound **Y**) followed by reaction with 3-bromopropionitrile in presence of potassium carbonate gave compound **19**. Selective protection of the primary alcohol with 4,4'-dimethoxytrityl chloride (DMTrCl) in pyridine gave **20**, which was then converted to its phosphoramidite **21** by reaction with bis(diisopropylamino) (2-cyanoethoxy)phosphine in the presence of tetrazole.



Scheme 5. Synthesis of the protected phosphoramidite of 2-methyl-6-thio-9-(2-deoxy-β-D-*erythro*-pentofuranos-1-yl)-purine. (a) NaH, 1-chloro-2-deoxy-3,5-di-*O*-p-toluoyl-α-D-*erythro*-pentofuranose, CH_3CN, rt, 2 h, 76 %; (b) $Na_2S_2O_3$, EtOH-H_2O, 120 °C, 12 h, 80%. (c) NaOMe, MeOH, 1.5 h, rt, 78%; (d) 3-bromopropionitrile, K_2CO_3, DMF, rt, 80%; (e) DMTrCl, pyridine, 0 °C to rt, 12 h, 75 %; (f) bis(diisopropylamino)(2-cyanoethoxy)phosphine, tetrazole, CH_2Cl_2, 0 °C to rt, 2.5 h, 68 %.

Hybridization properties of XNA modified DNA

The oligonucleotides were synthesized by using the phosphoramidite method and the reagents **14**, **17** and **21** on solid phase employing a DNA synthesizer. The synthesized sequences, together with the MS data, are shown in Table 1. The oligonucleotides ON 20-23 were used as well for T_m studies as for *in vivo* transliteration experiments.

Base pairing properties of the **X**, **Y** and **Z** bases were examined by hybridizing oligomers with their complementary strands and determining the T_m of the hybrids by temperature-dependent UV spectroscopy. The T_m was determined at 260 nm in NaCl (0.1M) buffer with KH_2PO_4 (20 mM, pH 7.5) and EDTA (0.1 mM) at a concentration of 4 µM for each strand. The stability of the duplexes containing the modified bases (**X**, **Y**, **Z**) was compared to the stability of the natural DNA duplex (Table 2).

10.1002/chem.201802304

Table 1. DNA sequences with **X**, **Y** and **Z** modified bases that have been synthesized, together with the MS data.

ON	Sequence	Mass	Mass
1	5'-CAC CGX TGC TAC C-3'	3907.7	3907.7
2	5'-GGT AGC AXC GGT G-3'	4067.7	4067.9
3	5'-CAC CGY TGC TAC C-3'	3908.7	3908.5
4	5'-GGT AGC AYC GGT G-3'	4068.7	4068.7
5	5'-GGT AGC AZC GGT G-3'	4012.7	4012.8
6	5'-d GCG CXX GCG C-3'	3096.6	3096.7
7	5'-GXG CAT GCX C-3'	3095.6	3095.6
8	5'-CTA GCX CCG TGC CAT GCA-3'	5472.0	5472.1
9	5'-p- TGC ATG GCA CGG XGC TAG-3'	5671.9	5671.8
10	5'-CTA GCX CXG TGC CAT GCA-3'	5526.0	5526.2
11	5'-p-TGC ATG GCA CXG XGC TAG-3'	5686.0	5685.9
12	5'-CTA GCX XCG TGC CAT GCA-3'	5526.0	5526.2
13	5'-p-TGC ATG GCA CGX XGC TAG-3'	5686.0	5685.9
14	5'-CTA GCX XXG TGC CAT GCA-3'	5580.0	5580.1
15	5'-p TGC ATG GCA CXX XGC TAG-3'	5700.0	5699.9
16	5'-p-CTA GCX CCG TGC CAT GCA-3'	5551,9	5552,2
17	5'-p-CTA GCX CXG TGC CAT GCA-3'	5605,9	5606,2
18	5'-p-CTA GCX XCG TGC CAT GCA-3'	5605,9	5606,2
19	5'-p-CTA GCX XXG TGC CAT GCA-3'	5660,0	5660,3
20	5'-p-CTA GCY CCG TGC CAT GCA-3'	5552.9	5552.8
21	5'-p-CTA GCY CYG TGC CAT GCA-3'	5607.9	5607.9
22	5'-p-CTA GCY YCG TGC CAT GCA-3'	5607.9	5608.0
23	5'-p-CTA GCY YYG TGC CAT GCA-3'	5662.9	5663.0
24	5'-p-CTA GCZ CCG TGC CAT GCA-3'	5496.9	5496.9
25	5'-p-CTA GCZ CZG TGC CAT GCA-3'	5495.9	5496.0
26	5'-p-CTA GCZ ZCG TGC CAT GCA-3'	5495.9	5496.2
27	5'-p-CTA GCZ ZZG TGC CAT GCA-3'	5494.9	5495.0
28	5'-p-TGC ATG GCA CGG ZGC TAG-3'	5616.9	5617.0
29	5'-p-TGC ATG GCA CZG ZGC TAG-3'	5575.9	5576.0
30	5'-p-TGC ATG GCA CGZ ZGC TAG-3'	5575.9	5576.1
31	5'-p-TGC ATG GCA CZZ ZGC TAG-3'	5534.9	5535.2
32	5'-AAC TGX GTC ATA GCT GTT TCC TG-3'	7059.5	7058.8
33	5'-AAC TGX XXG TCA TAG CTG TTT CCT G-3'	7745.7	7745.0

Table 2. T_m values (in °C) of antiparallel stranded DNA oligonucleotide duplexes A:, 5'-CAC CG**1** TGC TAC C-3' and B:, 3'-GTG GC**2** ACG ATG G-5' containing A, T, C, G bases and the modified base pairs **X**:**X**, **Y**:**Y**, **Y**:**Z** and **X**:**Z**.

Sequence A	**1**	**A**	**T**	**C**	**G**	**X**	**Y**	**Y**	**X**
Sequence B	**2**	**T**	**A**	**G**	**C**	**X**	**Y**	**Z**	**Z**
Duplex	$T_{m\,(°C)}$	57.5	57.3	61.3	59.7	45.2	41.7	47.3	46.6

From this table it is clear that the potential base pairing between the modified bases **X**:**X**, **Y**:**Y**, **Y**:**Z** and **X**:**Z** are very week, if they exist at all (certainly for **Y**:**Y**). For example, the pyrimidone nucleoside **Z** is not expected to occur in the conformation as proposed in Figure 1d. However, the stability **Z**:**X** and **Z**:**Y** is about the same. The T_m of the duplex is decreased with at least 10 °C when compared with the T_m of the canonical base pairs. This could be due to the difficulty of **X**, **Y** and **Z** to adopt a high energy *syn* or *anti* conformation within the duplex or due to steric hindrance. The T_m of duplexes with a modified pyrimidine, modified purine base pair (**Y**:**Z** and **X**:**Z**) is higher than the T_m of duplexes with a potential modified self-complementary purine base pair (**X**:**X** and **Y**:**Y**), respecting shape-recognition rules, and the difficulty for the simultaneous orientation of the **X** and **Y** modified purine bases in the *syn* and *anti* conformation. Steric clash between the hydrophobic substituent within the potential **X**:**X** and **Y**:**Y** base pair (despite the encouraging modeling data) may further explain these results.

In order to investigate the selectivity of base pairing, the stability of duplexes with modified **X**, **Y**, **Z** bases and the four canonical bases (A, T, G, C) was investigated by T_m determination (Table 3). In all cases, the potential base interactions are weak, especially with the **Z** base. The potential base pairing with the canonical bases is, in general, somewhat better than the potential **X**:**X** and **Y**:**Y** self-pairing. The highest T_m in this series is found for **Y**:**A**, which might be due to *syn*:*anti* interactions (Figure 6).

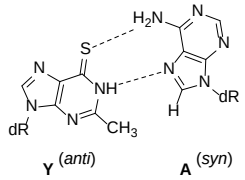


Figure 6. Potential base pairing between the 2-methyl-6-thiopurine base (**Y**) and adenine (**A**) in the **Y** (*anti*): **A** (*syn*) conformation.

As the strength of duplex formation is not only the result of potential interstrand base:base interactions by H-bonding but also due to potential base:base interactions by interstrand stacking, we have evaluated the **X**:**X** and the **Y**:**Z** system further by incorporating several modified **X, Y** and **Z** bases in a dsDNA, either separately or consecutively (Table 4). One of the sequences is phosphorylated at the 5'-position, because this sequence was also used for insertion in a plasmid for the *in vivo* experiments. The T_m data shows that multiple incorporations further reduced duplex stability in comparison to dsDNA reference duplex (T_m = 72.9 °C).

10.1002/chem.201802304

Table 3. T_m values (in °C) of antiparallel stranded DNA oligonucleotide duplexes containing A, T, C, G bases and the modified bases **X**, **Y**, **Z**.

Sequence	T_m (°C)	Sequence	T_m (°C)
5'-CAC CG**X** TGC TAC C-3'		5'-CAC CG**Y** TGC TAC C-3'	
3'-GTG GC**T** ACG ATG G-5'	49.1	3'-GTG GC**T** ACG ATG G-3'	50.8
3'-GTG GC**A** ACG ATG G-5'	48.2	3'-GTG GC**A** ACG ATG G-3'	51.1
3'-GTG GC**G** ACG ATG G-5'	47.7	3'-GTG GC**G** ACG ATG G-3'	47.8
3'-GTG GC**C** ACG ATG G-5'	47.2	3'-GTG GC**C** ACG ATG G-3'	44.7
5'-GGT AGC A**X**C GGT G-3'		5'-GGT AGC A**Z**C GGT G-3'	
3'-CCA TCG T**T**G CCA C-5'	49.3	3'-CCA TCG T**T**G CCA C-3'	43.7
3'-CCA TCG T**A**G CCA C-5'	49.0	3'-CCA TCG T**A**G CCA C-5'	45.6
3'-CCA TCG T**G**G CCA C-5'	48.9	3'-CCA TCG T**G**G CCA C-5'	48.3
3'-CCA TCG T**C**G CCA C-5'	49.4	3'-CCA TCG T**C**G CCA C-5'	42.0

Table 4. T_m values (in °C) of antiparallel stranded DNA oligonucleotide duplexes with single and multiple **X**, **Y** and **Z** incorporated.

Sequence	T_m (°C)	Sequence	T_m (°C)
5'-P-TGC ATG GCA CGG **X**GC TAG-3' 3'- ACG TAC CGT GCC **X**CG ATC-5'	61.7	5'-p-CTA GC**Y** CCG TGC CAT GCA-3' 3'-GAT CG**Z** GGC ACG GTA CGT-p-5'	63.2
5'-P-TGC ATG GCA C**X**G **X**GC TAG-3' 3'- ACG TAC CGT G**X**C **X**CG ATC-5'	51.3	5'-p-CTA GC**Y** C**Y**G TGC CAT GCA-3' 3'-GAT CG**Z** G**Z**C ACG GTA CGT-p-5'	54.0
5'-P-TGC ATG GCA CG**X** **X**GC TAG-3' 3'- ACG TAC CGT GC**X** **X**CG ATC-5'	54.7	5'-p-CTA GC**Y** **Y**CG TGC CAT GCA-3' 3'-GAT CG**Z** **Z**GC ACG GTA CGT-p-5'	55.7
5'-P-TGC ATG GCA C**XX** **X**GC TAG-3' 3'- ACG TAC CGT G**XX** **X**CG ATC-5'	49.0	5'-p-CTA GC**Y** **YY**G TGC CAT GCA-3' 3'-GAT CG**Z** **ZZ**C ACG GTA CGT-p-5'	48.8

Finally, we carried out three additional experiments to further analyze the properties of the 2-amino-6-methyl-8-oxo-7,8-dihydropurine (**X**) base: a) the potential of the **X** base to induce hairpin formation; b) a CD experiment to compare the shape of natural dsDNA with the "**X** incorporated" dsDNA; c) the potential of the triphosphate of the **X** nucleoside to function as substrate for DNA polymerases.

The 2-amino-6-methyl-8-oxo-7,8-dihydro-purine nucleosides **X** was incorporated in the self-complementary 5'-GCG CAT GCG-C-3' DNA sequence as well at the second position from both ends as in the middle positions and the T_m was measured at different concentrations ranging from 2 µM to 12 µM in 1N NaCl at 286 nm (Table 5). When positioning **X** at both ends, the oligonucleotide remains in a self-complementary duplex just as the natural sequence, although with considerably reduced T_m. However, positioning **XX** in the middle of the oligonucleotide, force it to adopt a hairpin conformation. This is in agreement with former observations that the **X**:**X** base pair is not really formed.

Table 5. Concentration dependent T_m of DNA after positioning **X** at different place.

Sequence	2 µM	4 µM	6 µM	8 µM	12 µM	Conformation
5'-GCG C**XX** GCG C-3'	64.9	64.6	64.5	64.1	64.5	Hairpin
5'-G**X**G CAT GC**X** C-3'	27.7	29.0	31.3	36.0	37.8	Duplex
5'-GCG CAT GCG C-3'	52.6	59.6	60.1	62.3	63.8	Duplex

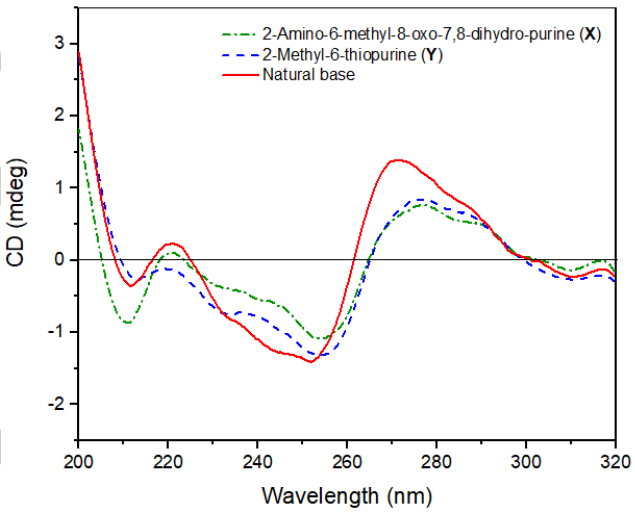


Figure 7. CD spectra of oligonucleotide duplexes. The ellipticity is plotted versus the wavelength for 2-amino-6-methyl-8-oxo-7,8-dihydro purine (dash dotted green line) and 2-methyl-6-thiopurine (dashed blue line) containing oligonucleotides along with control sequences with natural base (solid red line). Sequence: 5'-CACCG**X**/**Y**TGCTAC-C-3' and its complement I 5'-GCTAGCA**X**/**Y** CGGTG-3'. In the reference duplex, the base pair in the position of **X**/**Y** is A/T.

Figure 7 shows the CD spectra of dsDNA with a single **X**:**X** and **Y**:**Y** base pair incorporated in the middle of the sequence. All duplexes have CD spectra indicative of right-handed B-DNA, with small variation in peak position and ellipticity. The spectra for the natural DNA sequence exhibit broad negative peak at 252 nm and positive long wavelength peak at about 272 nm, typical of right handed B-DNA. The spectra for sequences with 2-amino-6-methyl-8-oxo-7,8-dihydro-purine **X** and 2-methyl-6-thiopurine **Y** resemble each other but show some deviation in peak position and ellipticity from the spectra of natural dsDNA.

10.1002/chem.201802304

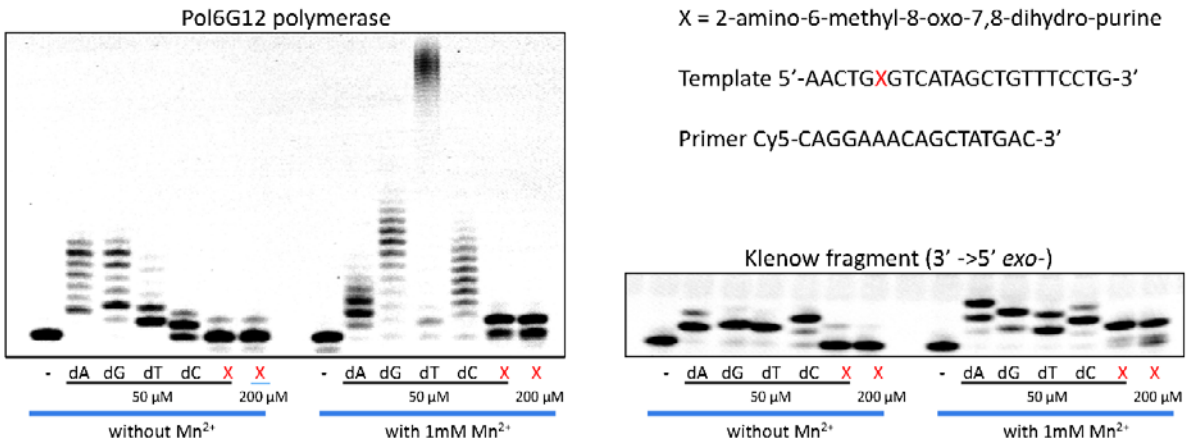


Figure 8. Polymerase catalysed incorporation of dNTP's and nucleotide **15** in a DNA primer-template complex.

The spectra shows a more narrow negative peak at 255 nm and positive peak at 276 nm with reduced height. The introduction of modified self–complementary base pair with hydrophobic groups not only destabilizes the DNA duplex but also results in small conformational changes.

The enzymatic incorporation of the modified **X** nucleotides in DNA was tested using Klenow (exo$^-$) and Pol6G12. The Klenow fragment was chosen because of its potential relevance for *in vivo* work in *E. coli*. The Pol6G12, an evolved enzyme derived from the thermophilic archaebacterium *Thermococcus gorgonarius*, was chosen because of its potential relevance for *in vitro* work (XNA gene synthesis and aptamer). First the triphosphate of the 2-amino-6-methyl-8-oxo-7,8-dihydro-9-(2-deoxy-β-D-*erythro*-pentofuranos-1-yl)-purine nucleoside was synthesized (**15**, Scheme 3). The incorporation of natural dNTP's and the modified nucleotide **X** was compared using a template with one and three nucleotides incorporated (see experimental section) and a DNA primer. Using Klenow polymerase in the absence of Mn^{2+} little **X** is incorporated against **X** in the template (Figure 8). Using Klenow polymerase in the presence of Mn^{2+}, the incorporation efficiency significantly increases (Figure 8). Data are similar for 50 µM and 200 µM of the modified nucleotide and using a template with one or three modified bases. In Figure 8, only data with one modified base in the template is shown. Similar results are observed using Pol6G12. In the presence of Mn^{2+}, the modified nucleotide **X** is incorporated against **X** in the template, although from previous experiments we know that **X**:**X** does not form a stable base pair, *in vitro*. We would also like to exclude that incorporation is due to a template-independent synthesis. In the presence of Mn^{2+} no template-independent synthesis is observed using Klenow polymerase (Figure 8). However, some template-independent synthesis is observed using Pol6G12 polymerase (although less than in the presence of the template).

Recognition of modified bases *in vivo*

In search for new orthogonal base pairs *in vivo*, it is important to know "if and how" the modified bases are functioning as template for canonical bases *in vivo* in the presence of cellular polymerases. To test the propagation of the XNA bases *in vivo*, the three bases were introduced separately (by chemical synthesis) in a sequence 5'-CTA **GCG CCG** TGC CAT GCA-3' representing the active site of thymidylate synthase (thyA) of *E. coli* (coding for Leu – Ala – Pro – Cys – His – Ala). They were introduced in the **GCG** codon (coding for Ala) and in the **CCG** codon (coding for Pro) (Table 6).

Table 6 Sequence of the synthesized oligonucleotides representing the active site of thyA of *E. coli* indicating the different positions (●) in which the modified bases (**X**, **Y**, **Z**) were introduced.

Code		Leu	Ala	Pro	Cys	His	Ala	
ON 16, 20, 24	5'-p-	CTA	GC●	CCG	TGC	CAT	GCA	-3'
ON 17, 21, 25	5'-p-	CTA	GC●	C●G	TGC	CAT	GCA	-3'
ON 18, 22, 26	5'-p-	CTA	GC●	●CG	TGC	CAT	GCA	-3'
ON 19, 23, 27	5'-p-	CTA	GC●	●●G	TGC	CAT	GCA	-3'

These oligonucleotides were ligated enzymatically in a gapped vector, transformed in a strain of *E. coli* (lacking thyA) and living colonies were analyzed (ratio between bacterial colony numbers in ± thymidine (dT) media). Only in case a productive DNA polymerase ternary complex is obtained with the unnatural nucleotide and a tolerated amino acid substitution is coded by the targeted codon, living colonies are obtained (as thyA in an essential enzyme in *E. coli*). Sequencing gives us information as which canonical base the modified bases recognize and which amino acids are inserted in the active protein. In each case, 20 clones were picked out in the dT deficient media and sequenced.

10.1002/chem.201802304

As shown in Figure 9, all three bases are very poorly recognized *in vivo*, when compared with the positive control. Even after introduction of only one base, the efficiency of transliteration of XNA to DNA is reduced to approximately 10 % in all cases. This percentage is further reduced between 0 % to 2 % when introducing two or three modified bases. When bars are not visible (Figure 9), only 1 or 2 colonies could be obtained, except for **Y** (2-methyl-6-thiopurine), inserted in GC(●) C(●)G, where no colony was obtained. These modified bases come close to the definition of orthogonal bases.

2-Amino-6-methyl-8-oxo-7,8-dihydro-purine, when introduced as third base in the Ala codon GCG (GC**X**) is recognized as a purine base (84 % as GCG and 16 % as GCA) (Table 7). Both triplets code for the amino acid "Ala". When 2-amino-6-methyl-8-oxo-7,8-dihydro-purine base is introduced as first or second base in the Pro codon (**X**CG and C**X**G), it is always (100 %) recognized as a cytosine base (CCG and CCG), coding for Pro (Table 8). This could be explained by X is recognized by G in a Wobble-Hoogsteen type association. These data are obtained using ON 17 and ON 18 where **X** is also incorporated in the Ala codon which might give some bias for base incorporation, especially for **X**CG (ON 18) It should be mentioned that not only CCG (coding for Pro) but also TCG (coding for Ser), GCG (coding for Ala) and CGG (coding for Arg) are tolerated at this position of the ThyA gene. However, we cannot conclude on the real impact of the individual amino acid substitutions on the enzyme efficiency. This observation is of interest as it shows that selection pressure (invariably Ala-Pro is selected) in this case is higher than structural constrains. The modified base can behave either as a purine or either as a pyrimidine analogue *in vivo* dependent on the place where it is inserted in the gene. Needless to say that there is no correlation with the results obtained by thermal stability experiment of the modified oligonucleotides (see previous), where hybridization of 2-amino-6-methyl-8-oxo-7,8-dihydro-purine **X** with A,G,C,T is very poor, if existing at all, and similar for all 4 canonical bases. Clearly, *E.coli* DNA polymerases have a key role in the selection procedure via stabilization of the protein-dsDNA-dNTP ternary complex.

The 2-methyl-6-thiopurine base **Y**, when incorporated in the GCG (Ala) codon at the third position (GC**Y**) is recognized almost exclusively (99 %) as a G (GCG) and in 1 % of the cases as a T (GCT) (Table 7). Both triplets code for Ala. When an extra modified base is introduced in the CCG (Pro) codon at the first position (**Y**CG), it is recognized for 93 % as G base (GCG) coding for Ala and for 7 % as C base (CCG) coding for Pro (Table 8). When 2-methyl-6-thiopurine base is incorporated in the second position of the same codon (C**Y**G) it is recognized as a cytosine base (CCG) in 100 % of the cases (Table 8). When introduced as well in the first as the second position (ON 23) of this CCG code (**YY**G), the codon only codes for Ala giving GCG as codon, and not anymore for Pro (data not shown in Table 8). It thus seems that the presence of one modified base in a codon may influence the recognition pattern of a second modified base, introduced in the same codon. Structural constrains seem more challenging than the selection pressure which leads to the high substitution of Pro by Ala (93 %). It should be remembered, of course, that only viable codons can be detected which is a bias of the analytical system (the obtained results are steered by the kind of amino acids that are accepted at a certain position in the protein). In the Ala codon (GC**Y**), however, this does not make a difference as all four possible codons (GCA, GCT, GCC, GCG) lead to the incorporation of Ala at this site of the protein (still a bias may exist based on the preferential use of *E. coli* for certain codons).

Finally, the 6-methyl-4-pyrimidone nucleoside **Z** was analysed *in vivo*, which gives a more complicated profile. In the Ala codon (GC**Z**), the 6-methyl-4-pyrimidone base is mainly recognized as a pyrimidine base (75 % as the T giving GCT, and 17 % as C giving GCC) and as G base to a minor extend (8 % as GCG) (Table 7). When introduced in the first position of the CCG code (**Z**CG) we find equal amounts of the TCG code (Ser) and CCG code (Pro) which means that 6-methyl-4-pyrimidone behaves as a pyrimidine base and the Ser substitution is well tolerated in thymidylate synthase (Table 8). The same amino acids (Ser and Pro) are coded for when the modified base is introduced at the first and second position (**ZZ**G). When introduced only at the second position (C**Z**G), the pyrimidone base is still recognized in majority as pyrimidine base (CCG), but also as G base in minority (CGG) coding for Arg. When three modified bases are introduced in a row (GC**Z** **ZZ**G), 58 % of Ala-Ser is introduced (GCT TCG) and 42 % of Ala-Pro (17 % GCG CCG, 17 % GCC CCG, 8 % GCT CCG) (data not shown in Table 8).

It is difficult to make general statements about the preferred recognition of modified purine and modified pyrimidine bases at this moment. Shape recognition (purine-pyrimidine base pair formation) is, however, important. The **Z** base is recognized as a pyrimidine in 90% of the cases (and only in 8-10% as a purine). Transversions coming from a purine-purine base pair are observed in relatively high yield with the **X** and **Y** base, probably because of the *syn-anti* equilibrium that may convert a purine to a pyrimidine-like base.

Table 7 - *In vivo* recognition of modified bases by canonical bases in the GCG (Ala) codon (aa: amino acid)

	5'-	CTA Leu	GC● aa	CCG Pro	TGC Cys	CAT His	GCA-3' Ala
	aa		**X** (%)		**Y** (%)		**Z** (%)
A	Ala		16		0		0
T	Ala		0		1		75
C	Ala		0		0		17
G	Ala		84		99		8

X 2-amino-6-methyl-8-oxo-7,8-dihydro-purine
Y 2-methyl-6-thio-purine
Z 6-methyl-4-pyrimidone

10.1002/chem.201802304

Table 8 - *In vivo* recognition of modified bases by canonical bases in the GCG (Pro) codon (aa: amino acid)

5'-GTA GC● ●CG TGC CAT GCA-3' Leu Ala aa Cys His Ala ●CG					5'-CTA GC● C●G TGC CAT GCA-3' Leu Ala aa Cys His Ala C●G			
	aa	**X** (%)	**Y** (%)	**Z** (%)	aa	**X** (%)	**Y** (%)	**Z** (%)
A	Thr	0	0	0	Glu	0	0	0
T	Ser	0	0	50	Leu	0	0	0
C	Pro	100	7	50	Pro	100	100	90
G	Ala	0	93	0	Arg	0	0	10

X 2-amino-6-methyl-8-oxo-7,8-dihydro-purine
Y 2-methyl-6-thio-purine
Z 6-methyl-pyrimidone

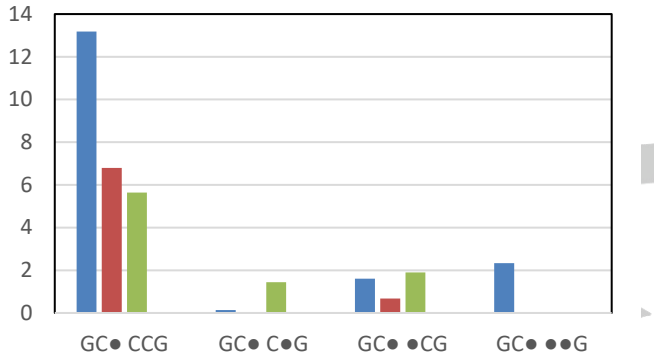


Figure 9. The *in vivo* propagation of the mosaïc DNA-XNA-DNA oligonucleotides. 2-amino-6-methyl-8-oxo-7,8-dihydro-purine modified oligonucleotides are represented in blue, 2-methyl-6-thio-purine modified oligonucleotides are represented in red and 6-methyl-4-pyrimidone modified oligonucleotides are represented in green. Ratios are normalized against positive control (100%) and correspond to the experimentally derived number of thymidine-prototrophic colonies (*thyA*+) from the total number of colonies.

Conclusions

An efficient synthetic scheme was developed for the synthesis of the protected phosphoramidites of 2'-deoxynucleosides with a 2-amino-6-methyl-8-oxo-7,8-dihydro-purine (**X**), a 2-methyl-6-thiopurine (**Y**) and a 6-methyl-4-pyrimidone (**Z**) base moiety. These building blocks were successfully used for oligonucleotide synthesis. These modified bases do not form base pairs with itself (**X**:**X** and **Y**:**Y**) and with the natural bases (A, G, C, T). Substitution of a polar functional group (NH_2, OH) of nucleobases by a methyl group at positions that are normally involved in H-bonding of canonical bases (the 6-carbonyl group of pyrimidine base, the 2-amino and 6-carbonyl group of purine base) readily leads to orthogonality (absence of base pairing with canonical bases). As demonstrated with the $\mathbf{T^H}$:$\mathbf{G^S}$ base pair[16] base pairing between methylated bases or bases with a thiocarbonyl group might easier occur when a H-atom is present in the opposite base (Figure 1c). However, the latter bases might not be really orthogonal. Surprisingly, some selected polymerases are still able to incorporate the 2-amino-6-methyl-8-oxo-7,8-dihydro-purine nucleoside opposite itself in a template. *In vivo*, 2-amino-6-methyl-8-oxo-7,8-dihydro-purine is mainly recognized as a A/G or C base. 2-Methyl-6-thio-purine is recognized as a G or C base (and in a minor case as T base). 6-Methyl-4-pyrimodone is *in vivo* mostly recognized as T, C but sometimes also as a G base. Although principles like shape recognition (purine-pyrimidine base pair formation) plays a role in the recognition process, it is not the only factor involved. The low efficiency of the transliteration process of XNA to DNA might implement some degree of randomness in the nucleotide selection procedure by the incorporated modified (XNA) nucleotide, when pyrimidine heterocycles (smaller heterocycles) are involved.

Experimental Section

^1H, ^{13}C and ^{31}P NMR spectra were recorded on 300 MHz (^1H NMR, 300 MHz; ^{13}C NMR, 75 MHz; ^{31}P NMR, 121 MHz) or 500 MHz (^1H NMR, 500 MHz; ^{13}C NMR, 125 MHz) or 600 MHz (^1H NMR, 600 MHz; ^{13}C NMR, 150 MHz) spectrometers. 2D NMRs (H-COSY, HSQC and HMBC) were used for the assignment of all the intermediates and final compounds. Mass spectra were acquired on a quadrupole orthogonal acceleration time-of-flight mass spectrometer. Column chromatographic separations were carried out by gradient elution with suitable combination of two/three solvents and silica gel (100–200 mesh or 230–400 mesh). Solvents for reactions were distilled prior to use: THF and toluene from Na/benzophenone; CH_2Cl_2 and CH_3CN from CaH_2; Et_3N and pyridine from KOH.

2-Chloro-4,5-diamino-6-methyl pyrimidine (**3**)**.** A solution of $Na_2S_2O_4$ (28.67 g, 164.6 mmol) in water (78 mL) was added to the stirred solution of **2** (10.35 g, 54.9 mmol) in THF (274 mL) and ethyl acetate (78 mL) at room temperature and the reaction was stirred for 3 hours. Ethyl acetate (200 mL) was added and extracted. The extract was dried over Na_2SO_4 and concentrated in vacuo. The crude mixture was purified by column chromatography using DCM:MeOH (9:1, v/v) as solvent to afford **3** (3.42 g, 39%). Analytical data of compound **3** was found to be identical to the reported compound.[25]

2-Chloro-6-methyl-8-oxo-7,8-dihydro-purine (**4**)**.** 1,1'-Carbonyl diimidazole (7.00 g, 43.2 mmol) was added to the solution of 2-chloro- 4,5-diamine-6-methyl-pyrimidine **3** (3.42 g, 21.6 mmol) in 1,4-dioxane (10 mL) and stirred at 110 °C under nitrogen for 50 min. Solvents were evaporated to obtain a yellow oil and CH_2Cl_2 (150 mL) was added. The flask shaken vigorously and allowed to settle the precipitate overnight. Solids were filtered through sintered funnel, washed with CH_2Cl_2 and dried under vacuum at 45 °C to afford **4** (3.29, 83%). ^1H NMR (300 MHz, DMSO-d_6) δ 12.00 (br s, 1H, NH), 11.39 (br s, 1H, NH), 2.37 (s, 3H, CH_3). ^{13}C NMR (75 MHz, DMSO-d_6) δ 153.7 (C6), 151.9 (C8),149.9 (C2), 144.2 (C4), 120.3 (C5), 18.7 (CH_3). HRMS (ESI+): calcd for $C_6H_5ClN_4O$ $[M + H]^+$ 185.0224; found 185.0219.

2-Chloro-6-methyl-8-oxo-7,8-dihydro-9-(2-deoxy-α/β-D-*erythro*-pentofuranos-1-yl)-purine (**6**). To a suspension of **4** (3.42 g. 18.5 mmol) in dry CH_3CN (142 mL) was added sodium hydride (60% in oil, 0.489 g,

20.4 mmol) and the mixture was stirred at room temperature under a nitrogen atmosphere for 30 min. 1-Chloro-2-deoxy-3,5-di-*O*-p-toluoyl-D-*erythro*-pentofuranose (7.20 g, 18.5 mmol) was added portion wise with stirring. The reaction mixture was stirred at 50 °C for 3 h and filtered to remove a small amount of insoluble material. Evaporation of the filtrate gave an oily residue, which was purified by silica gel column chromatography using ethyl acetate (35%) in hexane to afford nucleoside α- and β-**5** (3.00 g, 30%). HRMS (ESI+): calcd for $C_{27}H_{25}Cl_1N_4O_6$ [M + H]$^+$ 537.1535, found 537.1541.The anomeric mixture (1.50 g, 2.79 mmol) was treated with 5 M solution of sodium methoxide (1.05 mL) for 2 h at room temperature. The mixture was concentrated and the residue was purified by flash column chromatography (CH_2Cl_2/MeOH 8.5:1.5) to yield the anomeric mixture of compound α- and β-**6** as a white solid (0.798 mg, 95 %). HRMS (ESI+): calcd for $C_{11}H_{13}Cl_1N_4O_4$ [M + H]$^+$ 301.0698, found 301.0707.

2-Chloro-6-methyl-8-oxo-7,8-dihydro-9-(2-deoxy-5-*O*-dimethoxytrityl-β-D-*erythro*-pentofuranos-1-yl)-purine (7). Compound α- and β-**6** (320 mg, 1.41 mmol) was co-evaporated with dry pyridine twice under argon atmosphere and then dissolved in dry pyridine (15 mL). 4,4'–Dimethoxytrityl chloride (527 mg, 1.56 mmol) in CH_2Cl_2 (2 mL) was slowly added drop wise under argon atmosphere at 0 °C, then the mixture was stirred at room temperature for 2 h. The reaction was quenched by the addition of MeOH and the solvents were evaporated. The residue was dissolved in CH_2Cl_2 and washed with H_2O, the organic layers were dried on Na_2SO_4 and evaporated under argon atmosphere. The compound **7** was isolated as β anomer by column chromatography separation (CH_2Cl_2/MeOH/TEA 97:2:1) as a white foam (372 mg, 58%). ^{1}H and ^{13}C NMR Spectrum of compound **7** shows that the compounds contain small amounts of trimethylamine and pyridine as compound stabilizer used during synthesis and purification. These impurities were removed at the next step of synthesis. ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.87 (s, 1H, NH), 7.38 -7.33 (m, 2H, Ar), 7.22-7.13 (m, 7H, Ar), 6.74 (dd, *J* = 32.3, 9.0 Hz 4H, Ar), 6.18 (dd, *J* = 7.6, 5.35 Hz, 1H, H-1'), 5.31 (d, *J* = 3.0 Hz, OH), 4.45 (br S, 1H, 3'H), 3.96-3.93 (m, 1H, H-4'), 3.70 (s, 3H, OCH_3), 3.69 (s, 3H, OCH_3), 3.29 (dd, *J* = 10.1, 7.5 Hz, 1H, H-5a'), 3.11 (dd, *J* = 10.1, 3.8 Hz, 1H, H-5b'), 2.91-2.85 (m, 1H, H-2a'), 2.41 (s, 3H, CH_3), 2.19-2.14 (m, 1H, 2b'). ^{13}C NMR (125 MHz, DMSO-d_6) δ 148.0, 147.9, 142.2, 140.0, 139.7, 139.6 (C2, C4, C8, C6, Ar), 135.4, 135.1, 126.2, 125.8, 119.8, 119.7, 117.7, 117.6, 116.5, 113.9, 109.37, 102.9, 102.8 (C5, Ar), 75.9 (C4'), 71.1 (C1'), 61.1 (C3'), 54.6 (C-5'), 45.0, 44.9 (OCH_3), 26.1 (C2'), 8.9 (CH_3). HRMS (ESI+): calcd for $C_{32}H_{31}ClN_4O_6$ [M + H]$^+$ 603.2004; found 603.2009.

2-Chloro-6-methyl-8-oxo-7,8-dihydro-9-(2-deoxy-β-D-*erythro*-pentofuranos-1-yl)-purine (8). Compound **7** (0.431 mmol, 0.26 g) was dissolved in a mixture of CH_2Cl_2 and MeOH (1:1 *v/v*, 25 mL), and HCl in MeOH (0.14 mL of 1.25 mol L^{-1} solution) was added portion wise. The reaction was monitored by TLC CH_2Cl_2/MeOH(95:5). Once completed, the reaction mixture was neutralized with pyridine (1 mL), and the liquid was concentrated. The resulting oil was purified by column chromatography CH_2Cl_2/MeOH (96:4) to afford **8** (54 mg 70%). ^{1}H NMR (500 MHz, MeOD-d_4) δ 6.30 (dd, *J* = 7.5, 6.9 Hz, 1H, H1'), 4.59-4.57 (m, 1H Hz, H3'), 3.99-3.96 (m, 1H, H4'), 3.81 (dd, *J* = 12.1 Hz, 3.8, 1H, H5a'), 3.70 (dd, *J* = 12.0, 4.8 Hz, 1H, H5b'), 2.99 (ddd, *J* = 13.7, 7.5, 6.1 Hz, 1H, 2a'), 2,46 (s, 3H, CH_3), 2.20 (ddd, *J* = 13.5, 6.7, 3.0 Hz, 1H, H-2b'). ^{13}C NMR (125 MHz, DMSO-d_6) δ 152.7 (C8), 150.5 (C6), 150.2 (C4), 145.9 (C2), 119.3 (C5), 87.7 (C4'), 82.2 (C1'), 71.7 (C3'), 62.5 (C5'), 36.3 (C2'), 17.5 (CH_3). HRMS (ESI+): calcd for $C_{11}H_{13}ClN_4O_4$ [M + H]$^+$ 301.0698; found 301.0707.

2-N, N-dimethylformamidine-6-methyl-8-oxo-9-(2'-deoxy-α/β-D-*erythro*-pentofuranos-1-yl)-purine 12: To the solution of anomeric mixture of nucleoside **6** (1.05 gm, 3.49 mmol) in ethanol (10 mL) was added hydrazine (1.36 mL, 34.92 mmol, 80%). Reaction mixture was stirred at 84 °C for 12 hrs. The reaction mixture was concentrated and the crude unstable hydrazine derivative α- and β-**9** was used for next reaction immediately. To an ice-cold solution of crude compound α- and β-**9** (3.49 mmol) in acetic acid (23 mL) was added dropwise a solution of sodium nitrite (53.16 mol) in a least amount of water in an ice bath at −5 °C. The reaction mixture was allowed to stir at room temperature for 12 hrs and concentrated in vacuo to afford inseparable anomeric mixture of nucleoside α- and β-**10**. To the solution of crude nucleoside α- and β-**10** (3.49 mmol) in methanol was added PtO_2 (0.71 mmol) and reaction mixture was stirred under hydrogen atmosphere (1.38 bar) at 55 °C for 45 hrs. The reaction mixture was filtered through the celite and concentrated under vacuo to afford the anomeric mixture of 2-Amino-6-methyl-8-oxo-9-(2'-deoxy-D-ribofuranosyl)-purine α- and β-**11** (**X**), HRMS (ESI+): calcd for $C_{11}H_{15}N_5O_4$ [M + H]$^+$ 282.1196; found 282.1199. The crude nucleoside α- and β-**11** (0.860 mg, 3.06 mmol) was dissolved in methanol (30 mL) and dimethylformamide dimethyl acetal (15 ml) was added. Reaction mixture was stirred for 1 hour and concentrated on rotary evaporator to afford dimethylformamidine protected nucleoside diol α- and β-**12**, HRMS (ESI+): calcd for $C_{14}H_{20}N_6O_4$ [M + H]$^+$ 337.1618; found 337.1621.

2-N,N-dimethylformamidino-6-methyl-8-oxo-7,8-dihydro-9-(2-deoxy-5-*O*-dimethoxy trityl-β-D-*erythro*-pentofuranos-1-yl)-purine (13). Compound α- and β-**12** (444 mg, 1.32 mmol) was co-evaporated with dry pyridine twice under argon atmosphere and then dissolved in dry pyridine (26 mL). 4,4'–Dimethoxytrityl chloride (447 mg, 1.35 mmol) in CH_2Cl_2 (2 mL) was slowly added drop wise under argon atmosphere at 0 °C, then the mixture was stirred at room temperature for 3 h. The reaction was quenched by the addition of MeOH and the solvents were evaporated. The residue was dissolved in CH_2Cl_2 and washed with H_2O, the organic layers were dried on Na_2SO_4 and evaporated under argon atmosphere. The compound **13** was isolated as beta anomer by column chromatography (CH_2Cl_2/MeOH/Pyridine 96:3:1) as a white solid (360 mg, 43%). ^{1}H NMR (600 MHz, DMSO-d_6) δ 11.16 (s, 1H, NH), 8.29 (s, 1H, =CH-N), 7.32-7.31 (m, 2H, Ar), 7.20-7.14 (m, 7H, Ar), 6.77-6.70 (m, 4H, Ar), 6.20 (dd, *J* = 5.2, 7.74 Hz, 1H, H1'), 5.24 (d, *J* = 4.8 Hz,1H, OH), 4.49-4.45 (m, 1H, H3'), 3.89-3.86 (m, 1H, H4'), 3.70 (s, 3H, OCH_3), 3.69 (s, 3H, OCH_3), 3.25 (dd, *J* = 10.0, 7.5 Hz, 1H, H5a'), 3.09 (dd, *J* = 10.0, 3.54 Hz, 1H, H5b'), 3.00 (s, 3H, NCH_3), 3.01-2.97 (m, 1H, H2a'), 2.96 (s, 3H, NCH_3), 2.30 (s, 3H, CH_3), 2.11 (ddd, *J* = 13.2, 8.0, 5.4 Hz 1H, H2b'). ^{13}C NMR (151 MHz, DMSO-d_6) δ 160.4 (C2), 158.0, 157.9 (OCH_3), 156.9 (C=N), 152.6 (C8), 149.8 (C4), 145.2 (C6), 144.1, 135.8, 129.8, 129.6,127.8, 127.7, 126.5 (Ar), 114.9 (C5), 113.0, 112.9 (Ar), 85.5 (C-Ar_3), 85.3 (C4'), 80.5 (C1'), 71.2 (C3'), 64.8 (C5'), 55.1, 55.0 (OCH_3), 40.3 (C2'), 36.1, 36.4 (NCH_3), 19.1 (CH_3). HRMS (ESI+) calcd for $C_{35}H_{38}N_6O_6$ [M+H]+ 639.2925 found 639.2927.

2-N, N-dimethylformamidine-6-methyl-8-oxo-9-(2'-deoxy-β-D-*erythro*-pentofuranos-1-yl)-purine (β-12): The column of silica gel (10.0 g) was first prepared in dichloromethane containing 1% pyridine. This column was overlayed with a small buffer of sand followed by more silica gel (3.0 g) in dichloromethane containing 1% TFA. The compound **13** (100 mg) was then loaded in 3 mL of CH_2Cl_2. The column was eluted with gradient solvent system CH_2Cl_2/MeOH/Pyridine. The compound β-**12** was eluted with CH_2Cl_2/MeOH/Pyridine (96:3:1). ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.35 (s, 1H, NH), 8.54 (s, 1H, =CH-N), 6.18 (t, *J* = 7.4 Hz, 1H, H1'), 5.22 (d, *J* = 4.0 Hz,1H, OH), 4.89 (t, *J* = 7.4 Hz, 1H, OH), 4.42 (d, *J* = 2.7 Hz, 1H, H3'), 3.79 (dd, *J* = 7.4, 4.65 Hz, 1H, H4'), 3.60 (dt, *J* = 11.9, 4.0 Hz, 1H, H5b'), 3.49-3.44 (m, 1H, H5a'), 3.15 (s, 3H, NCH_3), 3.03 (s, 3H, NCH_3), 3.03-2.97 (m, 1H, H2a'), 2.34 (s, 3H, CH_3), 2.01 (ddd, *J* = 13.0, 6.7, 2.9 Hz, 1H, H2b'). ^{13}C NMR (125 MHz, DMSO-d_6) δ 158.4 (C=N), 154.3 (C6, C2), 151.8 (C8), 145.9 (C4), 117.1 (C5), 89.4 (C4'), 83.0 (C1'), 73.1 (C3'), 64.2 (C5'), 42.8 (NCH_3) 37.6 (C2'), 36.8 (NCH_3), 20.8 (CH_3). HRMS (ESI+) calcd for $C_{14}H_{20}N_6O_4$ [M+H]$^+$ 337.1618, found 337.1617.

3'-(2-cyanoethyl-N,N-diisopropyl)phosphoramidite of 2-N,N-dimethyl formamimidino-6-methyl-8-oxo-7,8-dihydro-9-(2-deoxy-5-*O*-

dimethoxytrityl-β-D-*erythro*-pentofuranos-1-yl)-purine (14). To a stirred solution of **13** (350 mg, 1.41 mmol) and 1 M solution of bis(diisopropylamino)(2-cyanoethoxy)phosphine (1.32 mL, 1.32 mmol) in anhydrous CH_2Cl_2 (10 mL), under argon atmosphere and at 0 °C, was added 0.45 M solution of 1H-terazole (1.62 mL, 0.728 mmol) dropwise. After 10 min, the ice bath was removed and reaction mixture was allowed to stir at room temperature for 45 min. The reaction mixture was diluted with CH_2Cl_2 and was washed with 1M TEAB solution. The extracts were dried over Na_2SO_4 and concentrated in vacuo. The crude mixture was purified by flash column chromatography (Hexane/Acetone/TEA, 80:20:1) to afford amidite **14** (300 mg, 62% yield). ^{31}P NMR (121 MHz, DMSO-d_6) δ 147.79, 147.41; HRMS (ESI+) calcd for $C_{44}H_{55}N_8O_7P_1$ [M+H]+ 839.4003, found 839.4009.

2-Amino-6-methyl-8-oxo-7,8-dihydro 8-9-(2-deoxy-5-*O*-triphosphate-β-D-*erythro*-pento-furanos-1-yl-purine (15). To an ice-cold solution of protected nucleoside β-**12** (30 mg, 0.09 mmol) in trimethyl phosphate (TMP) (1.0 mL) was added phosphoryl chloride (12 μL, 0.13 mmol.) and the solution stirred at 0°C for 5 hours. Tributylamine (105 μL, 0.44 mmol) and tetrabutylammonium pyrophosphate solution (240 mg, 0.27 mmol) in DMF (1 mL) was added simultaneously, and the solution stirred for a further 30 minutes. The reaction was then quenched by the addition of 0.5 M triethylammonium bicarbonate (TEAB) buffer (10 mL), and stored at 4°C overnight. The solvent was evaporated and the residue was treated with 25% ammonia (1 mL) for 15 min. The solution was evaporated to dryness and re-dissolved in water (5 mL) and applied to a Sephadex A25 column in 0.1 M TEAB buffer. The column was eluted with a linear gradient of 0.1-1.0 M TEAB. Appropriate fractions were pooled and evaporated to dryness to give desired product. The final purification was performed by RP HPLC (Alltima 5μ C-18 reverse phase column, buffer A, 0.1 M TEAB; buffer B, 0.1 M TEAB, 50% MeCN. 0% to 100% buffer B over 60 min). ^{31}P NMR (D_2O, δ, ppm): -6.38 (1P-γ, d, J 20 Hz), -11.91 (1P-α, d, J 20 Hz), -22.64 (1P-β, t, J 20 Hz). ESIHRMS found: *(M-H)*$^-$ 520.0041. calcd for $C_{11}H_{18}N_5O_{13}P_3$: (M-H)$^-$ 520.0041.

6-Methyl-3-(2-deoxy-β-D-*erythro*-pentofuranos-1-yl)-4(3*H*)-pyrimidone (**Z**). To a stirred solution of 6-methyl-4-pyrimidone (637 mg, 5.79 mmol) in acetonitrile (25 mL) at room temperature under argon was added bis(trimethylsilyl)acetamide (1.18 g, 5.79 mmol). After stirring at ambient temperature for 40 min, 1-chloro-2-deoxy-3,5-di-O-p-toluoyl-D-erythro-pentofuranose (1.5 g, 3.86 mmol) and dropwise $SnCl_4$ (3.87 mL, 3.86 mmol) was added. After 1 h, complete dissolution of the 2-deoxyribofuranoside had occurred. To the reaction was added EtOAc, and the resulting solution was successively extracted with saturated $NaHCO_3$ and brine. The organics were dried over anhydrous Na_2SO_4, and solvents were removed in vacuo. Purification via flash column chromatography on silica gel (30-50% ethyl acetate in hexanes) afforded a mixture of two anomers. The anomeric mixture was treated with 5 M solution of sodium methoxide (0.650 mL) for 2 h at room temperature. The mixture was concentrated and the residue was purified by flash column chromatography (CH_2Cl_2/MeOH 8.5:1.5) to yield the title compound **Z** as a white solid (315 mg, 36 %). ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.63 (s, 1H, H2), 6.25 – 6.11 (m, 2H, H5,H1'), 5.27 (d, *J* = 4.2 Hz, 1H, OH), 5.07 (t, *J* = 5.1 Hz, 1H, OH), 4.27 (dq, *J* = 7.5, 3.8 Hz, 1H, H3'), 3.87 (q, *J* = 3.6 Hz, 1H, H4'), 3.64 (ddd, *J* = 11.9, 5.0, 3.7 Hz, 1H, H5b'), 3.57 (ddd, *J* = 11.9, 4.7, 4.1 Hz, 1H, H5a'), 2.29 (ddd, *J* = 13.3, 6.2, 3.8 Hz, 1H, H2b'), 2.19 (s, 3H, CH_3), 2.12 – 2.08 (m, 1H, H2a'); ^{13}C NMR (151 MHz, DMSO-d_6) δ l163.4 (C4), 159.9 (C6), 147.7 (C2), 111.3 (C5), 88.0 (C4'), 84.4 (C1'), 70.1 (C3'), 60.9 (C5'), 41.1 (C2'), 23.1(CH_3); HRMS (ESI+) calcd for $C_{10}H_{14}N_2O_4$ [M+Na]$^+$ 249.0846, found 249.0847.

Further elution with (CH_2Cl_2/MeOH 8:2) afforded alpha isomer of compound **Z** as a white solid (157 mg, 18 %). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.41 (s, 1H, H2), 6.19 (s, 1H, H5), 6.16 (d, *J* = 5.9 Hz, 1H, H1'), 4.31 (t, *J* = 4.3 Hz, 1H, H4'), 4.25 (d, *J* = 5.5 Hz, 1H, H3'), 3.49 – 3.36 (m, 2H, H5), 2.61 – 2.53 (m, 1H, H2b'), 2.18 (s, 3H, CH_3), 1.95 (d, *J* = 14.2 Hz, 1H, H2a'); ^{13}C NMR (126 MHz, DMSO) δ 163.7 (C4), 160.4 (C6), 148.1 (C2), 111.6 (C5), 90.6 (C4'), 86.8 (C1'), 71.1 (C3'), 62.4 (C5'), 41.0 (C2'), 23.4 (CH_3); HRMS (ESI+) calcd for $C_{10}H_{14}N_2O_4$ [M+Na]$^+$ 249.08460, found 249.0853.

6-Methyl-3-(2-deoxy-5-*O*-dimethoxytrityl- β-D-*erythro*-pentofuranos-1-yl)-4(3*H*)-pyrimidone (16). Compound **Z** (320 mg, 1.41 mmol) was co-evaporated with dry pyridine (2 x 10 mL) under argon atmosphere and then dissolved in dry pyridine (15 mL). 4,4'–Dimethoxytrityl chloride (527 mg, 1.56 mmol) in CH_2Cl_2 (2 mL) was slowly added drop wise under argon atmosphere at 0 °C, then the mixture was stirred at room temperature for 2 h. The reaction was quenched by the addition of MeOH and the solvents were evaporated. The residue was dissolved in CH_2Cl_2 and washed with H_2O, the organic layers were dried on Na_2SO_4 and evaporated under argon atmosphere. Compound **16** was isolated by column chromatography (CH_2Cl_2/MeOH/TEA 97:2:1) as a white solid (520 mg, 69%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.40 (s, 1H, H2), 7.38 (d, *J* = 7.4 Hz, 2H, Ar), 7.33 – 7.19 (m, 7H, Ar), 6.88 (d, *J* = 8.4 Hz, 4H, Ar), 6.27 – 6.22 (m, 2H, H1', H5), 5.34 (d, *J* = 4.6 Hz, 1H, OH), 4.26 (dq, *J* = 9.0, 4.5 Hz, 1H, H3'), 3.99 (dd, *J* = 8.3, 4.5 Hz, 1H, H4'), 3.74 (s, 6H, 2CH_3), 3.24 (dd, *J* = 10.5, 5.5 Hz, 1H, H5'b), 3.20 (dd, *J* = 10.5, 3.1 Hz, 1H, H5'a), 2.34 (ddd, *J* = 13.4, 6.5, 4.5 Hz, 1H, H2b'), 2.22 (dd, *J* = 13.3, 6.6 Hz, 1H, H2a'), 2.18 (s, 3H, CH_3); ^{13}C NMR (126 MHz, DMSO-d_6) δ 163.5 (C4), 159.8, 158.1 (C6, Ar), 147.3 (C2), 144.8, 135.5, 135.4, 129.8, 127.9, 127.7, 126.8 (Ar), 113.3 (Ar), 111.6 (C5), 86.0 (C4'), 85.8 (C), 84.6 C1'), 70.2 (C3'), 63.6 (C5'), 55.0 (CH_3), 40.7 (C2'), 23.1 (CH_3); HRMS (ESI+) calcd for $C_{31}H_{32}N_2O_6$ [M+Na]$^+$ 551.2152, found 551.2151.

The 3'-(2-cyanoethyl-N,N-diisopropyl)phosphoramidite of 6-methyl-3'-(2-deoxy-5-*O*-dimethoxytrityl-β-D-*erythro*-pentofuranos-1-yl)-4(3H)-pyrimidone (17). To a stirred solution of **16** (350 mg, 1.41 mmol) and 1 M solution of bis(diisopropylamino)(2-cyanoethoxy)phosphine (1.32 mL, 1.32 mmol) in anhydrous CH_2Cl_2 (10 mL), under argon atmosphere and at 0 °C, was added 0.45 M solution of 1H-terazole (1.62 mL, 0.728 mmol) dropwise. After 10 min, the ice bath was removed and reaction mixture was allowed to stir at room temperature for 45 min. The reaction mixture was diluted with CH_2Cl_2 and was washed with 1M TEAB solution. The extracts were dried over Na_2SO_4 and concentrated in vacuo. The crude mixture was purified by flash column chromatography (Hexane/Acetone/TEA, 80:20:1) to afford amidite **17** (300 mg, 62% yield). ^{31}P NMR (121 MHz, DMSO-d_6) δ 147.75, 147.37; HRMS (ESI+) calcd for $C_{40}H_{49}N_4O_7P$ [M+Na]$^+$ 751.3231, found 751.3222.

2-Methyl-6-chloro-9-(2-deoxy-3,5-di-*O*-*p*-toluoyl-β-D-*erythro*-pentofuranos-1-yl)-purine (18a). A mixture of 6-chloro-2-methylpurine[26] (520 mg, 3.09 mmol) and 60% sodium hydride (164 mg, 4.11 mmol) in anhydrous acetonitrile (15 mL) was stirred at ambient temperature for 30 minutes under a nitrogen atmosphere. The 1-chloro-2-deoxy-3,5-di-*O*-p-toluoyl-D-*erythro*-pentofuranose (800 mg, 2.06 mmol) was added portionwise, and stirred for 2 hours. The reaction mixture was diluted with ethylacetate and washed with water. The organics were dried over anhydrous Na_2SO_4, and solvents were removed in vacuo. Purification via column chromatography on silica gel (30% ethyl acetate in hexanes) afforded the desired product **18a** as a white solid compound (815 mg, 76%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.41(s, 1H, H8), 7.96 (d, *J* = 8.0 Hz, 2H, Ar-H), 7.54 (d, *J* = 8.0 Hz, 2H, Ar-H), 7.29-7.25 (m, 2H, Ar-H), 7.20-7.16 (m, 2H, Ar-H), 6.64 (d, *J* = 6.6 Hz, 1H, H1'), 5.71 (d, *J* = 6.4 Hz, 1H, H3'), 4.91 (br S, 1H, H4'), 4.66-4.59 (m, 2H, H5'), 3.20 (d, *J* = 15.0 Hz, 1H, H2b'), 3.04 (dt, *J* = 15.0, 6.6 Hz, 1H, H2a'), 2.72 (s, 3H, CH_3), 2.43 (s, 3H, CH_3), 2.39 (s, 3H, CH_3); ^{13}C NMR (151 MHz, $CDCl_3$) δ 166.1, 165.7 (C=O), 162.5 (C2), 151.5 (C4), 150.4 (C6), 144.7, 144.3 (Ar), 142.5 (C8),

129.8, 129.6, 129.3, 129.3 (Ar, C5), 126.6, 125.7 (Ar), 86.3 (C1'), 84.7 (C4'), 74.8 (C3'), 63.9 (C5'), 38.3 (C2'), 25.6 (CH_3), 21.7 (CH_3), 21.6 (CH_3); HRMS (ESI+) calcd for $C_{27}H_{25}ClN_4O_5$ $[M+Na]^+$ 543.1405, found 543.1409

2-Methyl-6-sulfhydryl-9-(2-deoxy-3,5-di-*O*-*p*-toluoyl-β-D-*erythro*-pentofuranos-1-yl)-purine (18b): To a solution of compound **18a** (750 mg, 1.4 mmol) in 1:1-ethanol and water (40 mL), sodium thiosulfate (1.14 g, 7.2 mmol) was added and heated at 120 °C in a seal tube for 12 hours. The reaction mixture was concentrated and diluted with ethyl acetate and washed with water. The organics were dried over anhydrous Na_2SO_4, and solvents were removed in vacuo. Purification via column chromatography on silica gel (50% ethyl acetate in hexanes) afforded the desired product **18b** as a white solid compound (600 mg, 80%). ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.41 (s, 1H, H8), 7.93 (d, *J* =8.2 Hz, 2H, Ar-H), 7.81 (d, *J* =8.2 Hz, 2H, Ar-H), 7.36 (d, *J* =8.2 Hz, 2H, Ar-H), 7.29 (d, *J* =8.2 Hz, 2H, Ar-H), 6.46 (t, *J* =7.1 Hz, 1H, H1'), 5.80-5.76 (m, 1H, H3'), 4.64 (dd, *J* = 10.8, 4.2 Hz, H5b'), 4.58-4.52 (m, 2H, H5a', H4'), 3.21-3.16 (m, 1H, H2b'), 2.77 (ddd, *J* = 9.3, 6.7, 2.9 Hz, 1H, H2a'), 2.47 (s, 3H, CH_3), 2.39 (s, 3H, CH_3), 2.36 (s, 3H, CH_3); ^{13}C NMR (151 MHz, DMSO-d_6) δ 176.8 (C6), 165.5, 165.3 (C=O), 155.4 (C2), 144.2, 144.1, 143.9 (Ar-C, C4), 140.9 (C8), 133.9 (C5), 129.5, 129.4, , 129.3, 126.6, 126.5 (Ar-C), 83.8 (C1'), 81.8 (C4'), 74.8 (C3'), 63.9 (C5'), 35.8 (C2'), 21.3 (CH_3), 21.2 (CH_3), 21.0 (CH_3); HRMS (ESI+) calcd for $C_{27}H_{26}N_4O_5S$ $[M+H]^+$ 519.1696, found 519.1700.

2-Methyl-6-sulfhydryl-9-(2-deoxy-β-D-*erythro*-pentofuranos-1-yl)-purine (Y). 5 M solution of sodim methoxide (0.25 mL) solution was added to compound **18b** (650 mg, 1.25 mmol) and the mixture was stirred for 1.5 h at room temperature. The reaction mixture was concentrated and the residue was purified by flash column chromatography (CH_2Cl_2/MeOH 9:1) to yield the title compound **Y** as a white solid (276 mg, 78%). ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.41(s, 1H, H8), 6.28 (dd, *J* = 7.1, 4.0 Hz, 1H, H1'), 5.32 (d, *J* = 4.0 Hz, 1H, OH), 4.98 (t, *J* = 5.5 Hz, 1H, OH), 4.40-4.35 (m, 1H, H3'), 3.87- 3.83 (m, 1H, H4'), 3.59 (dt, *J* = 9.9, 4.8 Hz, 1H, H5a'), 3.53-3.48 (m, 1H, H5b'), 2.63- 2.56 (m, 1H, H2b'), 2.49 (s, 3H, CH_3), 2.28 (ddd, *J* = 9.9, 6.1, 3.3 Hz, 1H, H2a'); ^{13}C NMR (151 MHz, DMSO-d_6) δ 176.6 (C6), 155.3 (C2), 144.2 (C4), 140.7 (C8), 133.6 (C5), 88.1 (C1'), 83.5 (C4'), 70.7 (C3'), 61.6 (C5'), 48.7 (C2'), 21.0 (CH_3); HRMS (ESI+) calcd for $C_{11}H_{14}N_4O_3S$ $[M+H]^+$ 283.0859, found 283.0863.

2-Methyl-6-cyanoethylthio-9-(2-deoxy-β-D-*erythro*-pentofuranos-1-yl)-purine (19). To a solution of dry compound **Z** (25 mg, 0.88 mmol) in dry DMF (6 mL) was added powdered anhydrous potassium carbonate (474 mg, 3.54 mmol). This was followed by the addition of 3-bromopropionitrile (367 mg, 2.66 mmol) via a dry syringe. After stirring at room temperature for 12 hours the DMF was removed by coevaporation with xylenes and the residue was purified by flash column chromatography (CH_2Cl_2/MeOH 95:5) to yield the title compound **19** as a white solid (240 mg, 80%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.60 (s, 1H, H8), 6.41 (t, *J* = 7.1 Hz, 1H, H1'), 5.34 (d, *J* = 3.9 Hz, 1H, OH), 5.05 (t, *J* = 5.6 Hz, 1H, OH), 4.43 (br s, 1H, H3'), 3.89 (d, *J* = 2.7 Hz, 1H, H4'), 3.64-3.51 (m, 4H, SCH_2, H5'), 3.06-3.02 (t, *J* = 6.6 Hz, 2H, CH_2), 2.75-2.68 (m, 1H, H2a'), 2.64 (s, 3H, CH_3), 2.34-2.30 (m, 1H, H2b'); ^{13}C NMR (126 MHz, DMSO-d_6) δ 160.7, 157.5 (C2, C6), 148.9 (C4), 142.8 (C8), 129.3 (C5), 119.3 (CN), 88.1 (C4'), 83.8 (C1'), 70.8 (C3'), 61.7 (C5'), 39.5 (C2'), 25.6 (CH_3), 23.6 (CH_2S), 18.0 (CH_2CN); HRMS (ESI+) calcd for $C_{14}H_{17}N_5O_3S$ $[M+H]^+$ 336.1124, found 336.1134.

2-Methyl-6-cyanoethylthio-9-(2-deoxy-5-*O*-dimethoxytrityl-β-D-*erythro*-pentofuranos-1-yl)-purine (20). The reaction of compound **19** (220 mg, 0.62 mmol) in dry pyridine (13 mL) and 4,4'–dimethoxytrityl chloride (233 mg, 0.69 mmol) was carried out as described for compound **16** and purified by flash column chromatography (CH_2Cl_2/MeOH/TEA 97:2:1), affording compound **20** as a white solid (261 mg, 75%); ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.50 (s, 1H, H8), 7.31-7.29 (m, 2H, ArH), 7.24 – 7.13 (m, 7H, ArH), 6.79 – 6.72 (m, 4H, ArH), 6.43 (t, *J* = 6.4 Hz, 1H, H1'), 5.39 (d, *J* = 4.6 Hz, 1H, OH), 4.49-4.45 (m, 1H, H3'), 4.01 (dt, *J* = 6.8, 4.0 Hz, 1H, H4'), 3.71 (s, 3H, OCH_3), 3.70 (s, 3H, OCH_3), 3.60 (t, *J* = 6.8 Hz, 2H, SCH_2), 3.26 (dd, *J* = 10.1, 6.7 Hz, 1H, H5b'), 3.16 (dd, *J* = 10.1, 3.8 Hz, 1H, H5a'), 3.04 (t, *J* = 6.7 Hz, 2H, CH_2CN), 2.88 (dt, *J* = 12.8, 6.2 Hz, 1H, H2b'), 2.50 (s, 3H, CH_3), 2.35-2.37 (m, 1H, H2a'); ^{13}C NMR (151 MHz, DMSO-d_6) δ 160.7 (C2), 158.1, 158.0, 157.4 (C6, Ar), 148.9 (C4), 144.9 (Ar), 143.2 (C8), 135.6, 135.5 (Ar), 129.7, 129.6, 129.5, 127.7, 127.6, 126.6 (Ar, C5), 119.3 (CN), 113.1, 113.0 (Ar), 86.2 (C4'), 85.4 (C), 83.9 (C1'), 70.8 (C3'), 64.2 (C5'), 55.1, 55.0 (OCH_3), 38.4 (C2'), 25.6 (CH_3), 23.7 (CH_2), 18.0 (CH_2). HRMS (ESI+) calcd for $C_{35}H_{35}N_5O_5S$ $[M+H]^+$ 638.2431, found 638.2446.

The 3'-(2-cyanoethyl-*N*,*N*-diisopropyl)phosphoramidite of 2-methyl-6-cyanoethylthio-9-((2-deoxy-5-*O*-dimethoxytrityl-β-D-*erythro*-pentofuranos-1-yl)-purine (21). The reaction of compound **20** (250 mg, 0.39 mmol) and 1 M bis(diisopropylamino)(2-cyanoethoxy)phosphine (784 µl, 0.78 mmol) and 0.45 M 1H-terazole (958 µl, 0.431 mmol) in anhydrous CH_2Cl_2 (8 mL), was carried out as described for compound **21** and purified by flash column chromatography (CH_2Cl_2/MeOH/TEA 98:1:1) afforded amidite **21** (223 mg, 68% yield); ^{31}P NMR (202 MHz, DMSO-d_6) δ 147.86, 147.22; HRMS (ESI+) calcd for $C_{44}H_{52}N_7O_6PS$ $[M+H]^+$ 838.3509, found 838.3516.

Oligonucleotide synthesis.

Oligonucleotide assembly was performed with an Expedite DNA synthesizer (Applied Biosystems) by using the phosphoramidite approach. The oligomers were deprotected and cleaved from the solid support by treatment with aqueous ammonia (30%) for 1/2 h at 25 °C and 1 h at 55 °C. After gel filtration on a NAP-25 column (Sephadex G25-DNA grade from GE Healthcare) with water as eluent, the crude mixture was analyzed by using a Mono-Q HR 5/5 anion exchange column, after which purification was achieved by using a Mono-Q HR 10/100 GL column (Pharmacia) with the following gradient system: Buffer A = 10 mM $NaClO_4$ with 20 mM Tris-HCl in 15% CH_3CN, pH = 7,4. Buffer B = 600 mM $NaClO_4$ with 20 mM Tris-HCl in 15% CH_3CN, pH=7,4. The low-pressure liquid chromatography system consisted of a HITACHI Primaide organizer with a HITACHI Primaide 1410 UV detector and with a HITACHI Primaide 1110 pump and a Mono-Q HR 10/100 GL column (Pharmacia). The product-containing fraction was desalted on a NAP-25 column and lyophilised, followed by analysis by mass spectrometry.

UV melting experiments.

Oligomers were dissolved in a buffer solution containing NaCl (0.1 M), KH_2PO_4 (0.02 M, pH 7,5) and EDTA (0.1 mM). The concentration was determined by measuring the absorbance in Milli-Q water at 260 nm at 80°C. The following extinction coefficients were used: dA, ε=15.060; dC, ε=7.100; dG, ε=12.180; dT, ε=8.560; **X**, ε= 2089; **Y**, ε = 903; **Z**, ε = 2919. Extinction coefficient have been determined at 260 nm, which is not the maximum of absorption, but the wavelength at which the concentration of the oligonucleotide is determined. The concentration for each strand was 4 µM in all experiments. Melting curves were determined with a Varian Cary 100 BIO spectrophotometer. Cuvettes were maintained at constant temperature by water circulation through the cuvette holder. The temperature of the solution was measured with a thermistor that was directly immersed in the cuvette. Temperature control and data acquisition were carried out automatically with an IBM-compatible computer by using Cary WinUV thermal application software. A quick heating and cooling cycle was carried out to allow proper annealing of both strands. The samples were then heated from 10 to 80 °C at a rate of 0.2 °C min-1, and were cooled again at the same speed. Melting temperatures were

determined by plotting the first derivative of the absorbance as a function of temperature; data plotted were the average of two runs. Up and down curves showed identical Tm values.

Circular dichroism materials and methods.

For CD analysis, 4 μM of modified duplex [5'-d (CAC CG**X**/**Y** TGC TAC C)-3' and 5'-d (GGT AGC A**X**/**Y** C GGT G)-3')] and natural duplex [5'-d (CAC CGA TGC TAC C)-3' and 5'-d (GGT AGC ATC GGT G)-3')] were dissolved in buffer (0.1 M NaCl, 20 mM KH_2PO_4 and 0.1 mM EDTA) at pH 7. CD experiments were performed with Jasco J-1500 spectropolarimeter equipped with a Julabo F12 Peltier thermoelectric temperature control unit. Spectra were recorded at 25 °C using a 1 mm path length quartz cuvette with scanning from 320 to 200 nm. Three scans were averaged with a bandwidth of 1 nm and a 0.5 s averaging time at a rate of 50 nm/min.

Primer extension experiments by DNA polymerases.

The incorporation of natural dNTPs and the 2-amino-6-Methyl-8-oxo-7,8-dihydro-purine nucleoside triphosphate was carried out using a Cy5 fluorescence labeled primer and a template containing one or three 2-amino-6-methyl-8-oxo-7,8-dihydro-purine nucleosides, incorporated (ON 28, 29). The primer 5'Cy5-CAGGAAACAGCTATGAC-3' was annealed with one of the templates 5'-AACTG**X**GTCATAGCTGTTTCCTG-3' or 5'-AACTG**XXX**GTCATAGCTGTTTCCTG-3' in a 1:2 molar ratio by heating the mixture at 95 °C for 5 minutes followed by slow cooling to room temperature. A mesophilic DNA polymerase: Klenow fragment *exo-* (New England Biolabs) and a thermostable polymerase Pol6G12[27] were used. The final DNA polymerization mixtures each contained 125 nM primer-template duplex, reaction buffer (NEBuffer 2 was used for Klenow fragment *exo-* and ThermoPol buffer was used for Pol6G12) and 50 μM of the natural dNTPs or 50 μM or 200 μM of 2-amino-6-Methyl-8-oxo-7,8-dihydro-purine nucleoside triphosphate. Final concentration in a reaction of Klenow fragment *exo-* and Pol6G12 were 0.0017 U/μl and 0.0186 U/μl, respectively. One millimolar of $MnCl_2$ was added in the reaction as indicated. The reactions were performed at 37 °C for Klenow fragment *exo-* and 75 °C for Pol6G12. Aliquotes were taken after 60 min and quenched by adding to a double volume of quenching buffer (90 % formamide, 50 mM EDTA, and 0.05 % bromophenol blue) and heated at 95 °C for 5 min. The samples were separated on a 1 mm 15 % denaturing polyacrylamide gel and gel bands were visualized using the Typhoon FLA 9500 phosphor imager (GE Healthcare Life Sciences). The images were processed using ImageQuant TL v8.1.0.0 (GE Healthcare).

In vivo transliteration and mutational analysis directed by modified oligonucleotides ligated to plasmid DNA.

Oligonucleotides were dissolved in milliQ water to 1 mM and diluted ten fold before assay. Oligonucleotides were tested inside a gapped heteroduplex generated through the enzymatic modification of two different plasmids pAK1 and pAK2. The form of the heteroduplex is described elsewhere in details (S. Pochet, C. R. Biologies, 2003 and Pezo, Angew. Chem., 2013). [3a, 3c] A mix of equimolar (25 ng each) purified NheI and NsiI cut pAK1 and purified EcoRI cut and dephosphorylated pAK2 were resuspended in 10 mM Tris-HCl pH 7.5 with 100 mM NaCl in 20 μL final volume. The mixture was denaturated at 95 °C for 5 min before cooling to ambient temperature over 2 h for hybridization, before water dialysis through 0.05 μM nitrocellulose filter (Millipore) for 30 min. The oligonucleotides (100 pmoles), as well as a positive control (olignucleotide with intact catalytic residue codon) and two negative controls (water and oligomer with deleted catalytic residue codon) were added separately to the dialyzed heteroduplex mixture in 1x DNA ligase T4 reaction buffer (NEB) for 20 μL per sample. The mixture was then denaturated at 85 °C and cooled as before. Ligation was performed by adding 1 mM ATP and 5 U T4 DNA ligase (NEB) to the samples before overnight incubation at 16 °C. Ligated mixtures were dialyzed as before, and transformed by electroporation into fresh electro-competent *E. coli* K12 strain (Δ*thyA*::*aad*). Incubation of the electroporated bacteria was performed at 37 °C for 1 hour, before plating 100 μL of a serial dilution of the suspension onto Muller-Hinton (MH) media containing 100 μg mL^{-1} ampicillin (spreading the 10^0 and 10^{-1} dilutions) and onto the same media supplemented additionally with 0.3 mM thymidine (10^{-3} and 10^{-4} dilutions).

Acknowledgements

We thank Luc Baudemprez for technical assistance and Chantal Biernaux for editorial help. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement no. ERC-2012-ADG_20120216/320683, from FWO Vlaanderen, Belgium (G078014N) and from the KU Leuven Research Council (OT/1414/128). Mass spectrometry was made possible by the support of the Hercules Foundation of the Flemish Government (grant 20100225-7).

Keywords: methylated nucleobases • self-pairing • duplex stability • polymerases • transliteration

[1] D. A. Malyshev, K. Dhami, T. Lavergne, T. Chen, N. Dai, J. M. Foster, I. R. Corrêa, F. E. Romesberg, *Nature* **2014**, *509*, 385-388.

[2] Y. Zhang, J. L. Ptacin, E. C. Fischer, H. R. Aerni, C. E. Caffaro, K. San Jose, A. W. Feldman, C. R. Turner, F. E. Romesberg, *Nature* **2017**, *551*, 644-647.

[3] a) S. Pochet, P. A. Kaminski, A. Van Aerschot, P. Herdewijn, P. Marlière, *C. R. Biol.* **2003**, *326*, 1175-1184; b) A. T. Krueger, L. W. Peterson, J. Chelliserry, D. J. Kleinbaum, E. T. Kool, *J. Am. Chem. Soc.* **2011**, *133*, 18447-18451; c) V. Pezo, F. W. Liu, M. Abramov, M. Froeyen, P. Herdewijn, P. Marlière, *Angew. Chem. Int. Ed.* **2013**, *52*, 8139-8143; d) V. Pezo, G. Schepers, C. Lambertucci, P. Marlière, P. Herdewijn, *ChemBioChem* **2014**, *15*, 2255-2258; e) O. Bande, D. Braddick, S. Agnello, M. Jang, V. Pezo, G. Schepers, J. Rozenski, E. Lescrinier, P. Marlière, P. Herdewijn, *Chem. Sci.* **2016**, 7, 995-1010; f) O. Bande, R. Abu El Asrar, D. Braddick, S. Dumbre, V. Pezo, G. Schepers, V. B. Pinheiro, E. Lescrinier, P. Holliger, P. Marlière, P. Herdewijn, *Chem. Eur. J.* **2015**, *21*, 5009–5022.

[4] a) Z. Yang, F. Chen, J. B. Alvarado, S. A. Benner, *J. Am. Chem. Soc.* **2011**, *133*, 15105-15112; b) A. M. Sismour, *Nucleic Acids Res.* **2004**, *32*, 728-735; c) J. A. Piccirilli, S. A. Benner, T. Krauch, S. E. Moroney, S. A. Benner, *Nature* **1990**, *343*, 33-37.

[5] a) N. Tarashima, Y. Komatsu, K. Furukawa, N. Minakawa, *Chemistry - A European Journal* **2015**, *21*, 10688-10695; b) Y. Nomura, S. Kashiwagi, K. Sato, A. Matsuda, *Angew. Chem.* **2014**, *126*, 13058-13062.

[6] a) J. C. Delaney, J. Gao, H. Liu, N. Shrivastav, J. M. Essigmann, E. T. Kool, *Angew. Chem. Int. Ed.* **2009**, *48*, 4524-4527; b) J. Chelliserrykattil, H. Lu, A. H. F. Lee, E. T. Kool, *ChemBioChem* **2008**, *9*, 2976-2980.

[7] a) R. Yamashige, M. Kimoto, Y. Takezawa, A. Sato, T. Mitsui, S. Yokoyama, I. Hirao, *Nucleic Acids Res.* **2011**, *40*, 2793-2806; b) D. A. Malyshev, Y. J. Seo, P. Ordoukhanian, F. E. Romesberg, *J. Am. Chem. Soc.* **2009**, *131*, 14620-14621.

[8] B. D. Heuberger, D. Shin, C. Switzer, *Org. Lett.* **2008**, *10*, 1091-1094.

[9] I. Hirao, M. Kimoto, T. Mitsui, T. Fujiwara, R. Kawai, A. Sato, Y. Harada, S. Yokoyama, *Nat. Methods* **2006**, *3*, 729-735.

[10] a) F. Seela, R. Kröschel, *Org. Biomol. Chem.* **2003**, *1*, 3900-3908; b) D. Jiang, F. Seela, *J. Am. Chem. Soc.* **2010**, *132*, 4016-4024; c) C. Roberts, R. Bandaru, C. Switzer, *J. Am. Chem. Soc.* **1997**, *119*, 4640-4649; d) S. Hikishima, N. Minakawa, K. Kuramoto, Y. Fujisawa, M. Ogawa, A. Matsuda, *Angew. Chem. Int. Ed.* **2005**, *44*, 596-598.

[11] a) E. T. Kool, *Acc. Chem. Res.* **2002**, *35*, 936-943; b) M. Winnacker, E. T. Kool, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 12498-12508; c) S. A.

Benner, *Acc. Chem. Res.* **2004**, *37*, 784-797; d) I. Hirao, M. Kimoto, R. Yamashige, *Acc. Chem. Res.* **2012**, *45*, 2055-2065; e) A. W. Feldman, F. E. Romesberg, *Acc. Chem. Res.* **2018**, *51*, 394-403.

[12] a) G. T. Walker, *Nucleic Acids Res.* **1988**, *16*, 3091-3099; b) A. Ono, T. Ueda, *Nucleic Acids Res.* **1987**, *15*, 219-232.

[13] a) P. A. Limbach, P. F. Crain, J. A. McCloskey, *Nucleic Acids Res.* **1994**, *22*, 2183-2196; b) R. K. Kumar, D. R. Davis, *Nucleic Acids Res.* **1997**, *25*, 1272-1280; c) N. A. Siegfried, R. Kierzek, P. C. Bevilacqua, *J. Am. Chem. Soc.* **2010**, *132*, 5342-5344; d) A. T. Larsen, A. C. Fahrenbach, J. Sheng, J. Pian, J. W. Szostak, *Nucleic Acids Res.* **2015**, *43*, 7675-7687.

[14] H. P. Rappaport, *Nucleic Acids Res.* **1988**, *16*, 7253-7267.

[15] D. Lee, C. Switzer, *Nucleosides, Nucleotides and Nucleic Acids* **2015**, *34*, 424-432.

[16] D.-K. Lee, C. Switzer, *Bioorg. Med. Chem. Lett.* **2016**, *26*, 1177-1179.

[17] S. Pochet, P. Marliere, *C. R. Acad. Sci. III* **1996**, *319*, 1-7.

[18] a) K. E. McAuley-Hecht, G. A. Leonard, N. J. Gibson, J. B. Thomson, W. P. Watson, W. N. Hunter, T. Brown, *Biochemistry* **1994**, *33*, 10266-10270; b) M. Kouchakdjian, V. Bodepudi, S. Shibutani, M. Eisenberg, F. Johnson, A. P. Grollman, D. J. Patel, *Biochemistry* **1991**, *30*, 1403-1412.

[19] G. E. Plum, A. P. Grollman, F. Johnson, K. J. Breslauer, *Biochemistry* **1995**, *34*, 16148-16160.

[20] C. Cadena-Amaro, M. Delepierre, S. Pochet, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1069-1073.

[21] a) E. Sochacka, R. H., M. Cypryk, M. Sobczak, M. Janicka, K. Kraszewska, P. Bartos, A. Chwialkowska, B. Nawrot, *Nucleic Acids Research* **2015**, *43*, 2499-2512; b) S. M. Arif, K. Geethanandan, P. Mishra, A. Surolia, U. Varshney, M. Vijayan, *Acta Crystallographica Section D* **2015**, *71*, 1514-1527; c) R. P. Singh, A. Singh, H. V. Sirohi, A. K. Singh, P. Kaur, S. Sharma, T. P. Singh, *FEBS Open Bio* **2016**, *6*, 640-650; d) R. J. Zauhar, C. L. Colbert, R. S. Morgan, W. J. Welsh, *Biopolymers* **2000**, *53*, 233-248; e) J. Sheng, A. Larsen, B. D. Heuberger, J. C. Blain, J. W. Szostak, *J. Am. Chem. Soc* **2014**, *136*, 13916-13924; f) S. Yamada, T. Misono, S. Tsuzuki, *J. Am. Chem. Soc* **2004**, *126*, 9862-9872.

[22] J. Šponer, J. Leszczynski, P. Hobza, *The Journal of Physical Chemistry A* **1997**, *101*, 9489-9495.

[23] S. Gabriel, J. Colman, *Berichte der deutschen chemischen Gesellschaft* **1901**, *34*, 1234-1257.

[24] J. Ludwig, *Acta Biochim. Biophys. Hung.* **1981**, *16*, 131-133.

[25] R. N. Prasad, C. W. Noell, R. K. Robins, *J. Am. Chem. Soc.* **1959**, *81*, 193-197.

[26] A. L. Smith, N. D. D'Angelo, Y. Y. Bo, S. K. Booker, V. J. Cee, B. Herberich, F. T. Hong, C. L. Jackson, B. A. Lanman, L. Liu, N. Nishimura, L. H. Pettus, A. B. Reed, S. Tadesse, N. A. Tamayo, R. P. Wurz, K. Yang, K. L. Andrews, D. A. Whittington, J. D. McCarter, T. S. Miguel, L. Zalameda, J. Jiang, R. Subramanian, E. L. Mullady, S. Caenepeel, D. J. Freeman, L. Wang, N. Zhang, T. Wu, P. E. Hughes, M. H. Norman, *J. Med. Chem.* **2012**, *55*, 5188-5219.

[27] V. B. Pinheiro, A. I. Taylor, C. Cozens, M. Abramov, M. Renders, S. Zhang, J. C. Chaput, J. Wengel, S. Y. Peak-Chew, S. H. McLaughlin, P. Herdewijn, P. Holliger, *Science* **2012**, *336*, 341-344.

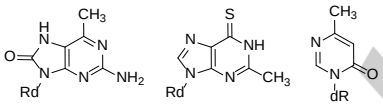
10.1002/chem.201802304

WILEY-VCH

Entry for the Table of Contents

FULL PAPER

Methylation of nucleobases at positions normally involved in W-C recognition of canonical bases readily leads to orthogonality.



A. M. Jabgunde,[a] F. Jaziri,[b] O. Bande,[a] M. Froeyen,[a] M. Abramov,[a] H. Nguyen,[a] G. Schepers,[a] E. Lescrinier,[a] V. B. Pinheiro,[c] V. Pezo,[b] P. Marlière,[b] and P. Herdewijn[a]*

Page No. – Page No.

Methylated nucleobases: Synthesis and evaluation for base pairing *in vitro* and *in vivo*