6-Substituted 2-Aminopurine-2′-deoxyribonucleoside 5′-Triphosphates that Trace Cytosine Methylation

Janina von Watzdorf and Andreas Marx*[a]

Introduction

Methylation of DNA in position 5 of cytosine (5mC) occurs predominantly at CpG islands[1] that are (largely) made up of CpG dinucleotides in promoter regions in the mammalian genome and is one of the most important epigenetic markers.[2] The role of 5mC in the mammalian genome is so distinct that this epigenetic marker is considered to be the “fifth base” of DNA.[3] 5mC serves as an important repressor of gene silencing,[4] thereby affecting tumorigenesis.[5] Cytosine 5-methylation is also the key mechanism mediating genomic imprinting[6] and silencing of foreign DNA[7] and is crucial for cellular development[8] and differentiation.[9]

Even though 5mC is the best-studied epigenetic modification,[10] some aspects are still poorly understood.[2] Detection of the occurrence and distribution of 5mC in the genome additionally holds the potential to serve as an important biomarker for diagnosis and disease therapy,[10] due to the important link of 5mC to human health.[11] Therefore, efficient methods for the detection of 5mC are required.

Different strategies for 5mC detection have already been described and used for genome-wide 5mC mapping. However, all of these methods—which rely variously on affinity enrichment,[12] endonuclease digestion,[13] nanopore sequencing,[14] and specific interactions of proteins with 5mC[15] or different chemical behaviour involving redox reactivity[16] or selective deamination with sodium bisulfite[17]—show several drawbacks. The most common method for the detection of 5mC is bisulfite sequencing, which can reveal the sites of epigenetic markers through comparison with the output of conventional sequencing methods.[18] Sodium bisulfite reversibly adds to the 5,6 double bond in pyrimidine nucleobases under acidic conditions, but no further reaction to replace the amino groups takes place.[19] Adjusting the pH to basic conditions converts non-5-methylated cytosine adducts into their uracil counterparts upon elimination.[20] In the case of 5mC, this deamination (to thymine) via the corresponding sulfonate adduct is nearly two orders of magnitude slower than that of the unmodified cytosine counterpart.[20] Bisulfite sequencing, taking advantage of this rate difference in the deamination of C and of 5mC, was subsequently established.

Although bisulfite sequencing has been used for genome-wide 5mC detection,[20] it has several drawbacks due to its reaction requirements.[21] The conditions used during bisulfite treatment are harsh and destroy around 95% of the genomic DNA[22] so large amounts of sample material are required. The method is time-consuming and prone to contamination[21] because many steps are needed and two sequencing runs are required for comparison. Additionally, deamination of C and 5mC after bisulfite treatment is deficient, leading to an error-prone output.[23]

Our goal, therefore, is a method for 5mC detection that can reveal the sites of epigenetic marks without the need to perform modification reactions prior to sequencing, thereby limiting the sources of error.

As reported previously[24] we discovered that 6-substituted 2-aminopurine-2′-deoxyribonucleoside 5′-triphosphates are able to sense 5mC in DNA polymerase-catalysed reactions. After testing several different purine-based 2′-deoxyribonucleotides with regard to their ability to be used in 5mC detection,
we found that modification in the 6-position was most promising. Nucleotides modified in this position were incorporated opposite C by Thermococcus kodakaraensis exo- DNA polymerase (KOD exo-) with notably different efficiencies than they were opposite 5mC, with favoured processing opposite C. In order to investigate whether the observed discrimination is more general and extendable to other modifications, we decided to explore different 6-substituted 2-aminopurine-2′-deoxyribonucleoside 5′-triphosphates that are modified at position 6 (Figure 1). We found that the observed effect is more general and not restricted to O′-alkyl modification of dGTP.

Results and Discussion

In order to investigate whether interruption of the Watson–Crick face of dGTP interferes with DNA polymerase-catalysed processing opposite C and 5mC, we decided to synthesise the nucleosides 2 and the secondary and tertiary amines 3 and 4 (Scheme 1).

The synthesis of the nucleotides 6-methylthio-dGTP (2a) and 6-ethylthio-dGTP (2b) is shown in Scheme 2. Because insertion of the modification by use of a nucleoside or nucleotide precursor failed, we introduced the modifications at the nucleobase level by nucleophilic replacement,[28] starting from commercially available 2-amino-6-chloropurine, as already reported for nucleotide 2a.[29] The modified nucleobases 5a and 5b were next used for glycosylation[27] to yield 6a and 6b. Subsequent deprotection of the hydroxy groups under basic conditions[26] resulted in nucleosides 7a and 7b, which were converted into the corresponding 5′-O-triphosphates 2a and 2b by standard triphosphorylation methods.[28]

To circumvent the need to perform laborious triphosphorylation reactions for every modified nucleotide, we established the synthesis of the 6-amino-modified dGTP derivatives 3a–g and 4a–h by treating the precursor nucleotide 2-amino-6-chloropurine-2′-deoxyriboside triphosphate (10) with aqueous solutions of the appropriate amines; yields ranged from 18–97 % (see Scheme 3). For the synthesis of the chlorinated nucleotide precursor 10, the hydroxy groups of commercially available 2′-deoxyguanosine required to be masked and afterwards selectively removed prior to 5′-triphosphorylation of the nucleoside. We employed tert-butyldimethylsilyl (TBDMs) groups and started the synthesis with 3′,5′-TBDMs-protected 2′-deoxyguanosine. By following known procedures,[29] we introduced chlorine in position 6 to obtain nucleoside 8 in a moderate yield of 37 %. The OH groups were subsequently deprotected by treatment with triethylamine trihydrofluoride in THF (93 %). Afterwards, 5′-triphosphorylation of 8 by standard procedures afforded 9.[27]

Next, the modified nucleotides were tested with regard to their action with 5mC. Thus, primer extension by single-nucleotide incorporation opposite C and 5mC followed by analysis by denaturing PAGE and visualisation by autoradiography were performed with the modified nucleotides in combination with three different DNA polymerases. The KlenTaq DNA polymerase, belonging to the sequence-family A,[30] and the KOD exo- DNA polymerase, a member of the sequence-family B,[30] were tested with regard to their ability to incorporate the modified nucleotides with increased differences in incorporation efficiencies opposite C or 5mC. Because previous experi-

Scheme 2. Synthesis of nucleotides 2a and 2b. a) 5a: Na2SMe, DMF, RT, 16 h (55 %); 5b: KOtBu, EtSH, DMF, reflux, 16 h (49 %); b) i: NaH (60 % in mineral oil), CH2CN, RT, 30 min; ii: 1-chloro-2-deoxy-3,5-di-O-toluoyl-α-ribofuranose, RT, 20 h (6a: 64 %, 6b: 59 %); c) 7 n ammonia in MeOH, 4 °C, 16 h (7a: 99 %, 7b: 61 %); d) i: NaX,Y,X′-tetramethylnaphthalene-1,8-diamine, trimethyl phosphate (TMP), POCl3, 0 °C, 30 min; ii: (Bu3NH)H3P2O5, DMF, nBu3N, RT, 30 min, 3. 0.1 M triethylammonium bicarbonate (TEAB), RT, 30 min (2a: 28 %, 2b: 28 %).
Figure 1. Discrimination observed on utilising 10 nM of A) KOD exo- DNA polymerase or B) 9’s North exo- DNA polymerase with indicated modified nucleotides (50 μM). Reactions were stopped after 5 min; % incorporation values opposite C or 5mC were calculated from the integrated gel band intensities [\% incorporation = (I_{\text{extension}} - I_{\text{primer}})/I_{\text{primer}}]. Discrimination was determined from the ratio of incorporation opposite C to incorporation opposite 5mC. Experiments were done in triplicate; arithmetic means are given. Values given in (A) for nucleotides 1a–d have already been published.\textsuperscript{[24]}

\[ I_{\text{primer}} = \text{Integrated gel band intensity of primer} \]
\[ I_{\text{extension}} = \text{Integrated gel band intensity of extension} \]
ments\textsuperscript{24} had shown the KOD exo- DNA polymerase to be most promising for the desired application, we tested the DNA polymerase 9’North exo- as a second sequence-family B polymerase as well.

As already indicated by previous experiments,\textsuperscript{24} no significant discrimination between C and 5mC was observed on employing KlenTaq DNA polymerase in combination with any of the modified nucleotides (Figure S1 in the Supporting Information). In addition, the incorporation efficiencies of KlenTaq DNA polymerase for the modified nucleotides decreased considerably with increasing size of the introduced modifications (Figure S1). Consequently, nucleotide 1a was incorporated with higher efficiencies than nucleotide 1d. The same holds true for the thio- and amino-modified nucleotides, with nucleotide 4a being processed with significantly higher efficiencies than nucleotide 4h bearing the bulky and inflexible pyrrolidine modification.

In contrast, on utilising the sequence-family B KOD exo- and 9’North exo- DNA polymerases we found that both DNA polymerases were capable of incorporating all modified nucleotides with only slightly decreased incorporation efficiencies relative to the natural dGMP. Interestingly, we observed a tendency towards more efficient processing of all modified nucleotides opposite C than opposite 5mC by both sequence-family B DNA polymerases (Figures 1, S2 and S4).

By employing KOD exo- DNA polymerase with the 6-thio-modified nucleotides 2, 2a and 2b we were able to observe enhanced discrimination for the commercially available nucleotide 6-thio-dGTP (2). This discrimination increased on introduction of a methythio (compound 2a) or ethylthio (compound 2b) group. Similar effects were observed for the nucleotides 6-methoxy-dGTP (1a) and 6-ethoxy-dGTP (1b).\textsuperscript{24} For the 6-amino-modified nucleotides 3a–g and 4a–h we found similar trends. Again, the discrimination increased with introduction of a methylamino (compound 3a) or ethylamino (compound 3b) group in comparison with 6-amino-dGTP (3). Interestingly, this effect vanishes after introduction of more bulky modifications (3c–g, Figures 1A and S2). Consequently, the decrease in discrimination due to increasing size of the modifications, such as in the introduction of a propylamino (compound 3c) or isopropylamino (compound 3d) group leads to discrimination ratios of only 2. On increasing the bulkiness of the modification even further through the introduction of a cyclopentyl group (compound 3e), the discrimination ratio decreased to less than 2. This discrimination ratio could be slightly increased again by employing the 3-azidopropanamino group (compound 3f). However, the discrimination ratio still does not surpass the discrimination observed on incorporation of 3b.

In addition, the difference in incorporation opposite C and 5mC could be further increased by employment of tertiary amines (compounds 4a–h). Discrimination observed with use of the nucleotide modified with the doubly methylated amino group (compound 4a) is higher than the discrimination observed for incorporation of 3a. Consistent with those trends observed above, discrimination increased with increasing size of the introduced modifications up to a certain degree (compound 4c), to decrease again when the steric hindrance of the employed modifications was enhanced further (compounds 4d–g). Surprisingly, though, the highest discrimination could be observed with the most rigid, sterically hindered pyrrolidine modification (compound 4h). In this case, a threefold higher incorporation opposite C could be observed than opposite 5mC (Figures 1, 2B and S2).

As already reported,\textsuperscript{24} we observed decreased selectivity when employing the dGTP derivatives because most of the modified nucleotides were incorporated opposite T with high efficiency as well (Figure S3).

Similar discrimination between C and 5mC was also observed with employment of 9’North exo- DNA polymerase. Consistent with the studies using the KOD exo- DNA polymerase, higher incorporation rates opposite C than opposite 5mC could already be seen when using the unmodified dGTP (Figure 1B). However, in the case of the 9’North exo- DNA polymerase, this discrimination could be enhanced only slightly by the introduction of alkoxy modifications at position 6 of the dGTP (compounds 1a–d). The discrimination increased slightly with growing bulkiness from the methoxy-modified nucleotide 1a to the propoxy-modified nucleotide 1c, but de-
increased again for the isopropoxy-modified nucleotide 1d. The replacement of the oxygen at position 6 by sulfur in nucleotides 2, 2a and 2b led to decreased discrimination for processing of those nucleotides opposite C and 5mC, in comparison with dGTP and the modified nucleotides 1a-d. However, we were able to enhance the discrimination for the 9’ North exo-DNA polymerase by utilising the 6-amino-modified nucleotides 3, 3a-g and 4a-h (Figures 1B and S4). Enhanced discrimination was already observable with an amino group (compound 3), but introduction of bulkier modifications such as methylamino (compound 3a), ethylamino (compound 3b) and propylamino (compound 3c) groups resulted in higher discrimination, whereas the bulkier modifications of nucleotides 3d-g led to decreased discrimination (Figures 1B and S4). Surprisingly, the discrimination could again be further enhanced by employment of the tertiary amine compounds 4a-h. Increasing the size of the introduced modification from nucleotide 4a to 4b leads to a twofold increase in discrimination. Nevertheless, a decline in discrimination ratios can be observed with employment of nucleotides 4c-f, bearing sterically more demanding modifications. Surprisingly, the highest discrimination, by a factor of 3, can be observed with use of nucleotide 4g with the sterically most demanding modifications (Figures 1B, 2C and S4).

Conclusion

In summary, we have synthesised various 6-substituted 2-aminopurine-2’-deoxyribonucleoside 5’-triphosphates and tested them with regard to their ability to discriminate between C and 5mC in the template strand. All modified nucleotides were accepted and processed by all tested DNA polymerases. In accordance with our previous findings, employing the KlenTaq DNA-polymerase failed to show any promising differences in incorporation efficiencies opposite C or opposite 5mC. In contrast, sequence-family B enzymes KOD exo- and 9’ North exo- DNA polymerases delivered results that are more promising. The different discrimination behaviour of the A-family and B-family DNA polymerases during processing of the modified nucleotides is striking. Structural data for both enzyme families bound to a primer/template complex show significantly different interaction patterns of the respective protein with its substrate. This difference might be responsible for the observed differences in incorporation behaviour.

We found increased discrimination between C and 5mC, to a certain degree, on enhancing the bulkiness of the introduced modification. Interestingly, discrimination could be further enhanced by the introduction of even bulkier tertiary amine groups. Our results indicate that sequence-family B DNA polymerases are best suited for the desired purpose. We identified the best discrimination between C and 5mC for KOD exo- DNA polymerase with the nucleotide 6-pyrolidine-dGTP (4h) and for 9’ North exo- DNA polymerase with nucleotide 6-(ethyl-isopropylamino)-dGTP (4g). We thereby verified the results of previous studies that used the 6-alkoxy-modified dGTP derivatives 1a-d. Interestingly, the degree of discrimination is not exclusively dependent on the bulkiness of the modification. It appears that a subtle interplay between the modification and the DNA polymerase is important for achieving the highest discrimination.

Consistently with our previous studies, we observed decreased selectivity of the modified nucleotides between C and T. This surely hinders its application in whole sequencing approaches. However, microarray-based approaches that address single CpG sites and have been reported for methylation profiling after bisulfite-mediated conversion might benefit from the findings reported here. The DNA methylation arrays offer low-cost alternatives for profiling large numbers of samples.

This study strongly suggests that sensing of 5-methylation is a more general phenomenon of this class of nucleotides and not restricted to O6-alkyl modification of 2-aminopurine-2’-deoxyribonucleoside 5’-triphosphates because it applies to other functionalities as well.
We have developed an extended toolbox to examine incorporation efficiencies opposite different nucleotide analogues in RNA and DNA; its biological impact is currently being studied intensively.2,5 We see great potential in applying this approach to the detection of different modifications in nucleic acids. We are convinced that this approach will help in providing more detailed understanding of the underlying principles resulting in discrimination between C and 5mC by employment of modified nucleotides.

**Experimental Section**

2-Amino-6-ethyliothiopurine (5b).58) Potassium tert-butoxide (2.7 g, 24.0 mmol, 10.0 equiv) was dissolved in abs. DMF (20 mL). Ethane-thiol (1.7 g, 24.0 mmol, 1 equiv) was added, and the reaction mixture was stirred at room temperature for 30 min. 2-Amino-6-chloropurine (0.4 g, 2.4 mmol, 1.0 equiv) was then added, and the reaction mixture was heated at reflux in a sealed vessel. After 16 h, the suspension was concentrated in vacuo, and the remaining solid was dissolved in water (10 mL). The aqueous solution was neutralised with acetic acid and extracted with ethyl acetate. The organic layer was dried over MgSO4 and concentrated to dryness, and the crude residue was purified by flash column chromatography with ethyl acetate as solvent to yield the product (0.23 g, 1.17 mmol, 61 %).

15.3, 152.1, 139.1, 124.3, 22.4, 15.6 ppm; HNMR (400 MHz, [D8]DMSO): δ = 2.24 (L, 1H), 7.87 (s, 1H), 6.29 (brs, 2H), 3.25 (q, J = 6.7 Hz, 2H), 1.32 ppm (t, J = 7.3 Hz, 3H); 13C NMR (100 MHz, [D8]DMSO): δ = 160.1, 159.3, 152.1, 139.1, 124.3, 22.4, 15.6 ppm; HNMR (400 MHz, [D8]DMSO): δ = 2.18 (L, 1H), 7.87 (s, 1H), 6.29 (brs, 2H), 3.25 (q, J = 6.7 Hz, 2H), 1.32 ppm (t, J = 7.3 Hz, 3H); 13C NMR (100 MHz, [D8]DMSO): δ = 160.1, 154.0, 149.0, 140.3, 124.0, 87.4, 83.1, 72.3, 63.0, 38.9, 26.2, 26.1, 18.4, 18.2, 4.43, 4.50, 5.1 ppm; HNMR (ESI): m/z calc for C16H14CN5O2S: 514.2431 [M+H]+; found: 514.2399.

6-Chloro-2'-deoxyguanosine (9): 3',5'-Di-O-tert-butylmethylsilyl-6-chloro-2'-deoxyguanosine (1.0 g, 1.9 mmol, 1 equiv) was dissolved in abs. THF (10 mL) at room temperature. Triethylamine trihydrofluoride (196 µL, 1.2 mmol, 1.7 equiv) was added and the reaction mixture was stirred for 16 h. The solution was then concentrated in vacuo and the crude residue was purified by flash column chromatography with methylene chloride with up to 10 % methanol to afford the product (504.0 mg, 1.7 mmol, 93 %).

1H NMR (400 MHz, [D8]DMSO): δ = 8.34 (s, 1H), 6.94 (brs, 2H), 6.22 (t, J = 6.7 Hz, 1H), 4.92 (t, J = 5.8 Hz, 1H), 4.37 (q, J = 3.5, 6.7 Hz, 1H), 3.83 (dt, J = 2.9, 4.6 Hz, 1H), 3.58 (dt, J = 5.1, 11.7 Hz, 1H), 2.61 (ddd, J = 5.8, 7.4, 13.2 Hz, 1H), 2.26 (ppm (ddd, J = 3.4, 6.2, 13.2 Hz, 1H); 13C NMR (100 MHz, [D8]DMSO): δ = 160.1, 154.1, 149.9, 140.5, 124.0, 87.4, 83.1, 72.3, 63.0, 38.9, 26.2, 26.1, 18.4, 18.2, 4.43, 4.50, 5.1 ppm; HNMR (ESI): m/z calc for C16H14CN5O2S: 514.2431 [M+H]+; found: 514.2399.

**General procedure A**26) Typical reaction scales ranged from 40 to 120 mg of starting nucleoside. The appropriate nucleoside (1.0 equiv) and proton sponge (N,N,N'-tetramethylthiphelenalene-1,8-diamine, 1.5 equiv) were dried in vacuo, dissolved in dry triethylphosphate (1 mL per 20 mg of starting nucleoside) at room temperature and cooled to 0 °C. Phosphorous oxychloride (12 equiv) was added dropwise at 0 °C and the mixture was stirred under nitrogen. After 30 min TLC showed complete conversion of starting material, and a solution (0.5 M) of (BuNH)2HPO4 in anhydrous DMF (5.0 equiv) and nBu4N (10.0 equiv) were added simultaneously to the mixture. The mixture was allowed to warm to room temperature and stirred for 30 min. Aqueous triethylamine bis-carbonate buffer (pH 7.5, 0.1 M, 10 mL) was then added and the mixture was stirred for a further 30 min. The aqueous layer was washed with ethyl acetate several times to remove trimethylphosphate and then concentrated to dryness. The residue was dissolved...
in water and purified by ion-exchange chromatography (DEAE Se- phadex A-25, buffer A: TEAB (0.1 M), buffer B: TEAB (1 M), linear gra- dient: 0% B to 100% B) and further purified by reversed-phase (RP) HPLC (Nucleodur RP 18-HTec, buffer A: triethylammonium ace- tate (TEAA, 50 mM), buffer B: acetonitrile, linear gradient: 5% B to 100% B). The triphosphates were concentrated to dryness. To get

6-Thiothioyl-2'-deoxyguanosine-5'-O-triphosphate (2b): Yield: 28% (90.0 μmol). 1H NMR (400 MHz, D2O): δ = 8.22 (s, 1H), 6.31 (t, J = 6.9 Hz, 1H), 4.72–4.63 (m, 1H), 4.26–4.21 (m, 1H), 4.18–4.10 (m, 2H), 3.26–3.17 (m, 2H), 2.81 (dt, J = 6.8, 13.8 Hz, 1H), 2.55 (ddd, J = 3.3, 6.3, 14.2 Hz, 1H), 1.33 ppm (t, J = 7.4 Hz, 1H). 13C NMR (100 MHz, D2O): δ = 158.3, 155.7, 145.7, 135.6, 120.3, 82.0, 79.5, 67.5, 61.9, 34.8, 19.4, 10.1 ppm; 13P NMR (162 MHz, D2O): δ = 10.22 to 11.1 (m, 1P), 11.1–12.0 (m, 1P), 20.5–22.5 ppm (m, 1P); HRMS (ESI): m/z calc for C13H11N2O5P2: 549.9958 [M–H]–; found: 549.9957.

6-Chloro-2'-deoxyguanosine-5'-O-triphosphate (10): Yield: 13% (178.0 μmol). 1H NMR (400 MHz, D2O): δ = 8.35 (s, 1H), 6.32 (t, J = 6.8 Hz, 1H), 4.78–4.75 (m, 1H), 4.25–4.22 (m, 1H), 4.18–4.15 (m, 2H), 2.84 (dd, J = 6.2, 7.5, 13.7 Hz, 1H), 2.55 ppm (ddd, J = 3.6, 6.4, 14.0 Hz, 1H); 13C NMR (100 MHz, D2O): δ = 159.4, 152.8, 150.2, 142.1, 123.5, 85.3, 87.3, 70.8, 65.5, 38.1 ppm; 13P NMR (162 MHz, D2O): δ = 10.86 (d, J = 20.1 Hz, 1P), –11.53 (d, J = 20.3 Hz, 1P), –23.5 ppm (t, J = 20.2 Hz, 1P); HRMS (ESI): m/z calc for C16H13ClN2O5P2: 523.9535 [M–H]–; found: 523.9525.

6-Methylthio-2'-deoxyguanosine-5'-O-triphosphate (3a): Yield: 93% (93.3 μmol). 1H NMR (400 MHz, D2O): δ = 8.09 (s, 1H), 6.27 (t, J = 6.9 Hz, 1H), 4.77 (dq, J = 3.3, 6.3 Hz, 1H), 4.26–4.23 (m, 1H), 4.19–4.12 (m, 2H), 3.02 (brs, 3H), 2.76 (ddd, J = 6.1, 7.6, 13.8 Hz, 1H), 1.99 ppm (brs, 3H); 13C NMR (100 MHz, D2O): δ = 159.6, 154.9, 149.1, 136.5, 112.8, 85.4, 82.9, 71.0, 65.5, 38.7, 23.0 ppm; 13P NMR (162 MHz, D2O): δ = 9.72 (d, J = 19.3 Hz, 1P), –11.28 (d, J = 19.3 Hz, 1P), –22.84 ppm (t, J = 19.8 Hz, 1P); HRMS (ESI): m/z calc for C16H14N2O5P: 519.0201 [M–H]–; found: 519.0189.

6-(Ethylamino)-2'-deaminopurine-2'-deoxyribonucleoside-5'-O-triphosphate (4a): Yield: 85% (85.3 μmol). 1H NMR (400 MHz, D2O): δ = 8.07 (s, 1H), 6.27 (t, J = 6.9 Hz, 1H), 4.70–4.65 (m, 1H), 4.27–4.14 (m, 3H), 3.30 (brs, 6H), 2.74 (dt, J = 6.9, 13.8 Hz, 1H), 2.53 ppm (ddd, J = 3.3, 6.2, 13.9 Hz, 1H); 13C NMR (100 MHz, D2O): δ = 157.3, 152.0, 153.9, 112.9, 85.5, 83.2, 71.0, 65.5, 38.8 ppm; 13P NMR (162 MHz, D2O): δ = 10.59 (d, J = 19.9 Hz, 1P), –11.46 (d, J = 19.9 Hz, 1P), –23.25 ppm (t, J = 19.8 Hz, 1P); HRMS (ESI): m/z calc for C16H14N2O5P: 533.0358 [M–H]–; found: 533.0344.

6-Methylthio-2'-deaminopurine-2'-deoxyribonucleoside-5'-O-triphosphate (4b): Yield: 73% (73.3 μmol). 1H NMR (400 MHz, D2O): δ = 8.03 (s, 1H), 6.24 (dd, J = 6.2, 7.5 Hz, 1H), 4.73 (dt, J = 3.5, 6.4 Hz, 1H), 4.23–4.19 (m, 1H), 4.18–4.10 (m, 2H), 3.75–3.64 (m, 2H), 3.09–3.05 (m, 3H), 2.64 (ddd, J = 6.1, 7.7, 13.7 Hz, 1H), 2.42 (ddd, J = 3.3, 6.2, 13.9 Hz, 1H), 1.05 ppm (t, J = 7.1 Hz, 3H); 13C NMR (100 MHz, D2O): δ = 158.5, 153.3, 150.8, 112.8, 88.5, 82.9, 71.1, 65.5, 45.6, 38.7, 35.9, 12.1 ppm; 13P NMR (162 MHz, D2O): δ = –10.10 (m, 1P), –11.42 (d, J = 19.5 Hz, 1P), –22.88 ppm (m, 1P); HRMS (ESI): m/z calc for C16H14N2O5P: 547.0514 [M–H]–; found: 547.0584.
6-Methylpropylamino-2-aminopurine-2'-deoxyribonucleoside 5'-O-triphosphate (4c): Yield: 67% (6.7 μmol). 1H NMR (400 MHz, D2O): δ = 8.15 (s, 1 H), 6.40–6.36 (m, 1 H), 4.74–4.68 (m, 1 H), 4.32–4.17 (m, 3 H), 3.94 (brs, 2 H), 3.36 (brs, 3 H), 2.82–2.83 (m, 1 H), 2.57–2.53 (m, 1 H), 1.73 (d, J = 7.5 Hz, 2 H), 0.92 ppm (t, J = 7.3 Hz, 3 H); 13C NMR (162 MHz, D2O): δ = 8.21; 11.39 (2P), –22.10 to –23.33 (m, 3 H). HRMS (ESI): m/z calcd for C13H19NO5P3: 561.0671 [M+H]+; found: 561.0681.

6-Methyl-isopropylamino-2-aminopurine-2'-deoxyribonucleoside 5'-O-triphosphate (4d): Yield: 34% (3.4 μmol). 1H NMR (400 MHz, CD3OD): δ = 8.18 (s, 1 H), 6.49 (t, J = 6.8 Hz, 1 H), 4.87 (dt, J = 3.4, 6.4 Hz, 1H), 4.42–4.40 (m, 2H), 4.39–4.26 (m, 1H), 3.49 (p, J = 1.6 Hz, 1.33), 1.67 (brs, 3 H), 2.93 (dt, J = 6.7, 13.5 Hz, 1 H), 2.54 (dd, J = 3.6, 6.3, 13.4 Hz, 1H), 1.41 ppm (t, J = 6.8 Hz, 6H); 13C NMR (162 MHz, CD3OD): δ = −2.78 (d, J = −12.6 Hz, 1 H), 1.11 (d, J = 20.4 Hz, 1P), −23.71 ppm (t, J = 21.2 Hz); HRMS (ESI): m/z calcd for C13H19NO5P3: 561.0671 [M+H]+; found: 561.0710.

6-Diethylamino-2-aminopurine-2'-deoxyribonucleoside 5'-O-triphosphate (4e): Yield: 65% (6.5 μmol). 1H NMR (400 MHz, D2O): δ = 8.27 (s, 1 H), 6.27 (dd, J = 6.2, 7.6 Hz, 1H), 4.75 (dqs, J = 2.3, 3.3 Hz, 1H), 4.28–4.10 (m, 3H), 3.79–3.69 (m, 4H), 2.76–2.69 (m, 1H), 2.50 (dd, J = 3.4, 6.3, 13.9 Hz, 1H), 1.16 (t, J = 7.0 Hz, 6H); 13C NMR (100 MHz, D2O): δ = 158.7, 153.0, 150.8, 135.5, 112.8, 85.4, 83.0, 79.0, 65.4, 43.2, 38.7, 16.7 ppm; 31P NMR (162 MHz, D2O): δ = −10.02 (m, 1P), −11.40 (d, J = 19.6 Hz, 1P), −22.89 ppm (t, J = 19.7 Hz); HRMS (ESI): m/z calcd: 561.0671 [M+H]+; found: 561.0673.

6-Ethylpropylamino-2-aminopurine-2'-deoxyribonucleoside 5'-O-triphosphate (4f): Yield: 48% (4.8 μmol). 1H NMR (400 MHz, CD3OD): δ = 7.97 (s, 1 H), 6.29 (t, J = 6.8 Hz, 1H), 4.69 (dt, J = 3.3, 6.2 Hz, 1H), 4.24–4.14 (m, 2H), 4.11–4.08 (m, 1H), 3.89 (brs, 2H), 3.82 (brs, 2H), 2.73 (dt, J = 6.4, 13.6 Hz, 1H), 2.34 (dd, J = 3.5, 6.2, 13.5 Hz, 1H), 1.69 (d, J = 7.4 Hz, 2H), 1.21 (t, J = 7.0 Hz, 3H), 1.03 ppm (t, J = 7.4 Hz, 3H); 31P NMR (162 MHz, CD3OD): δ = −0.10 (d, J = 21.3 Hz, 1P), −11.19 (d, J = 20.5 Hz, 1P), −23.64 ppm (t, J = 21.5 Hz); HRMS (ESI): m/z calcd for C13H19NO5P3: 575.0827 [M+H]+; found: 575.0835.

6-Ethyl-isopropylamino-2-aminopurine-2'-deoxyribonucleoside 5'-O-triphosphate (4g): Yield: 27% (2.7 μmol). 1H NMR (400 MHz, D2O): δ = 8.07 (s, 1H), 6.29 (t, J = 6.8 Hz, 1H), 5.27 (m, 1H), 4.75 (m, 1H), 4.24–4.11 (m, 3H), 3.60 (m, 2H), 2.71 (dt, J = 6.8, 13.7 Hz, 1H), 2.49 (dd, J = 3.3, 6.4, 14.2 Hz, 1H), 1.17 (d, J = 6.8 Hz, 6H), 1.13 ppm (t, J = 6.3 Hz, 3H); 31P NMR (162 MHz, D2O): δ = −10.27 (m, 1P), −11.14 (m, 1P), −22.93 ppm (m, 1P); HRMS (ESI): m/z calcd for C13H19NO5P3: 575.0827 [M+H]+; found: 575.0826.

6-Pyrrolidine-2-aminopurine-2'-deoxyribonucleoside 5'-O-triphosphate (4h): Yield: 18% (1.8 μmol). 1H NMR (400 MHz, D2O): δ = 8.16 (s, 1H), 6.36 (t, J = 6.7 Hz, 1H), 4.86–4.80 (m, 1H), 4.31–4.27 (m, 1H), 4.25–4.17 (m, 2H), 4.07–3.55 (m, 4H), 2.80 (dt, J = 6.8 Hz, 13.8 Hz, 1H), 2.54 (dd, J = 3.4, 6.3, 14.0 Hz, 1H), 2.11–2.01 ppm (m, 4H); 31P NMR (162 MHz, D2O): δ = −10.26 (m, 1P), −11.26 (d, J = 19.9 Hz, 1P), −23.11 ppm (t, J = 18.6 Hz); HRMS (ESI): m/z calcd for C13H19NO5P3: 559.0514 [M+H]+; found: 559.0541.

Acknowledgements

We gratefully acknowledge the European Research Council (ERC Advanced Grant 339834) and the Konstanza Research School Chemical Biology for financial support.

Keywords: SmC - DNA methyltransferase - DNA polymerases - epigenetics - gene expression - modified nucleotides
[30] J. Ito, D. K. Braithwaite, Nucleic Acids Res. 1991, 19, 4045–4057.
[31] K. Bergen, K. Betz, W. Welte, K. Diederichs, A. Marx, ChemBioChem 2013, 14, 1058–1062.
[32] a) J. Sandoval, H. Heyn, S. Moran, J. Serra-Musach, M. A. Pujana, M. Bibikova, M. Esteller, Epigenetics 2011, 6, 692–702; b) M. Bibikova, B. Barnes, C. Tsan, V. Ho, B. Klotzle, J. M. Le, D. Delano, L. Zhang, G. P. Schroth, K. L. Gunderson, J. B. Fan, R. Shen, Genomics 2011, 98, 288–295; c) S. Dedeurwaerder, M. Defrance, E. Calonne, H. Denis, C. Sotiriou, F. Fuks, Epigenomics 2011, 3, 771–784.
[33] a) Y. Fu, C. He, Curr. Opin. Chem. Biol. 2012, 16, 516–524; b) I. A. Roundtree, C. He, Curr. Opin. Chem. Biol. 2016, 30, 46–51; c) L. Zhang, K. E. Szulwach, G. C. Hon, C. X. Song, B. Park, M. Yu, X. Lu, Q. Dai, X. Wang, C. R. Street, H. Tan, J. H. Min, B. Ren, P. Jin, C. He, Nat. Commun. 2013, 4, 1517; d) X. Wang, C. He, Mol. cell 2014, 56, 5–12; e) B. S. Zhao, C. He, Cell Res. 2015, 25, 153–154.

Manuscript received: April 26, 2016
Accepted article published: June 2, 2016
Final article published: July 6, 2016