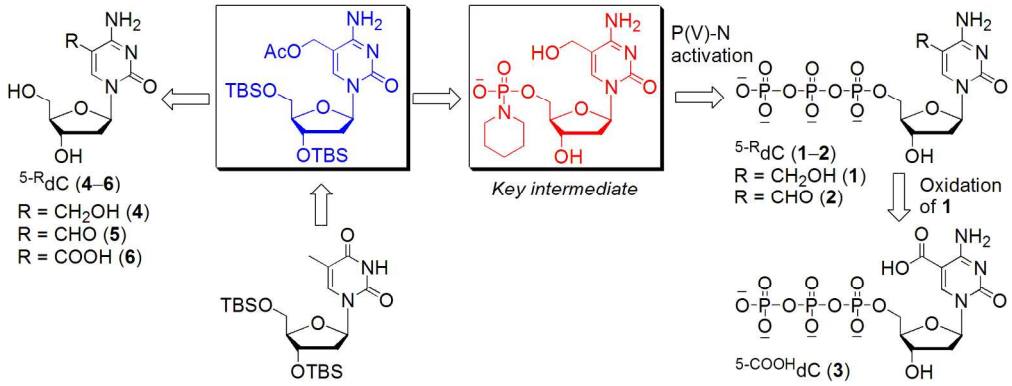
RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances



180x68mm (300 x 300 DPI)

RSC Advances

RSCPublishing

COMMUNICATION

Efficient synthesis of 5-hydroxymethyl-, 5-formyl-, and 5-carboxyl-2′-deoxycytidine and their triphosphates

Cite this: DOI: 10.1039/x0xx00000x

Qi Sun,*[a] Jian Sun,[a] Shan-Shan Gong,[a] Cheng-Jun Wang,[a] Shou-Zhi Pu[a] and Fu-De Feng*[b]

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

An efficient P(V)–N activation strategy for the preparation of high-quality 5-hydroxymethyl-, 5-formyl-, and 5-carboxyl-2′-deoxycytidine triphosphates has been developed. The method was also optimized for gram-scale synthesis of the corresponding parent nucleosides from 2′-deoxythymidine.

During the past few decades, it has been unravelled that DNA methyltransferases-mediated methylation of cytosine in eukaryotic genomes is one of the most important epigenetic marks for transcriptional gene silencing.[1] In mammalian DNA, cytosines are predominantly methylated within CpG sites, and the methylation patterns are heritable as stable epigenetic signals during cell divisions.[2] Meanwhile, DNA demethylation in specific contexts is required for the recovery of cytosine bases, enabling flexible and dynamic regulation of gene expression during cellular development.[3]

Recently, a series of 5-methylcytosine (5-mC) oxidation products, 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC), and 5-carboxycytosine (5-caC) have been detected in mammalian DNA.[4] The ten eleven translocation proteins (TET1–3) were identified as the corresponding oxidases with molecular oxygen and 2-ketoglutarate as cofactors. The new research evidence reported by He et al.[6] and Zhang et al.[7] strongly supported that TET-mediated oxidation of 5-mC leads to active DNA demethylation in epigenetic programming of cells.[8] While thymine-DNA glycosylase (TDG) was determined as an enzyme for 5-caC excision repair,[6] other enzymes involved in either base excision repair (BER) or decarboxylation pathways remain to be elucidated.

In the past few years, several phosphoramidite-based solid phase approaches for the preparation of 5-hmC-, 5-fC-, and 5-caC-containing oligodeoxynucleotides (ODNs) have been developed.[9] To advance the investigation of the mechanisms and enzymes related to 5-mC oxidation, longer DNA fragments with 5-hmC-, 5-fC-, and 5-caC bases are highly desired but hard to be synthesized by the tedious solid phase methods. More recently, Carell and his co-workers reported an expeditious synthesis of long 5-hmC-, 5-fC-, and 5-caC-containing ODNs (150 bp) from 5-hydroxylmethyl-, 5-formyl-, and 5-carboxyl-2′-deoxycytidine triphosphates ($^{5\text{-HOMe}}$dCTP (**1**), $^{5\text{-CHO}}$dCTP (**2**), and $^{5\text{-COOH}}$dCTP (**3**)) by polymerase chain reaction (PCR).[10] However, the oxidative modifications of cytosine posed a huge challenge for the preparation of the corresponding triphosphates (Fig. 1). Though the reported yield for $^{5\text{-CHO}}$dCTP was 70%, the disproportional peaks on its ^{31}P NMR spectrum revealed that the sample was contaminated with a significant amount of polyphosphate impurities.[10] Similar issues were also found in the low-yielding synthesis of $^{5\text{-COOH}}$dCTP (7% yield) and the protected $^{5\text{-HOMe}}$dCTP (1% yield).[10] In this paper, we report an efficient preparation of the triphosphates of all three $^{5\text{-Me}}$dC oxidation products (**1**–**3**) on the basis of the P(V)–N activation strategy we established for nucleoside polyphosphate synthesis.[11] The optimized method for gram scale synthesis of the parent nucleosides, $^{5\text{-HOMe}}$dC (**4**), $^{5\text{-CHO}}$dC (**5**), and $^{5\text{-COOH}}$dC (**6**), from dT is also described.

Currently, there are two major synthetic routes for the preparation of $^{5\text{-Me}}$dC oxidation derivatives. The one utilizing 5-iodo-2′-deoxycytidine ($^{5\text{-I}}$dC) starting material directly installed the 5-formyl or 5-methoxycarbonyl group by the Pd-catalyzed Stille reaction.[9a–d] The reduction of $^{5\text{-CHO}}$dC (**5**) afforded $^{5\text{-HOMe}}$dC (**4**).[9a,10] However, the cost of $^{5\text{-I}}$dC and the use of pressurized reactor for gaseous CO limited its application. Therefore, we employed the other approach starting from dT.[4c]

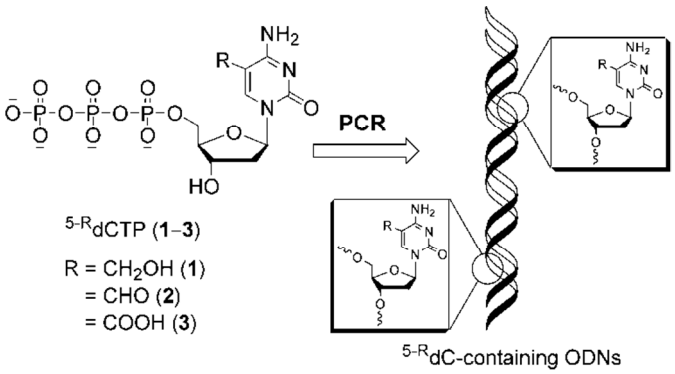
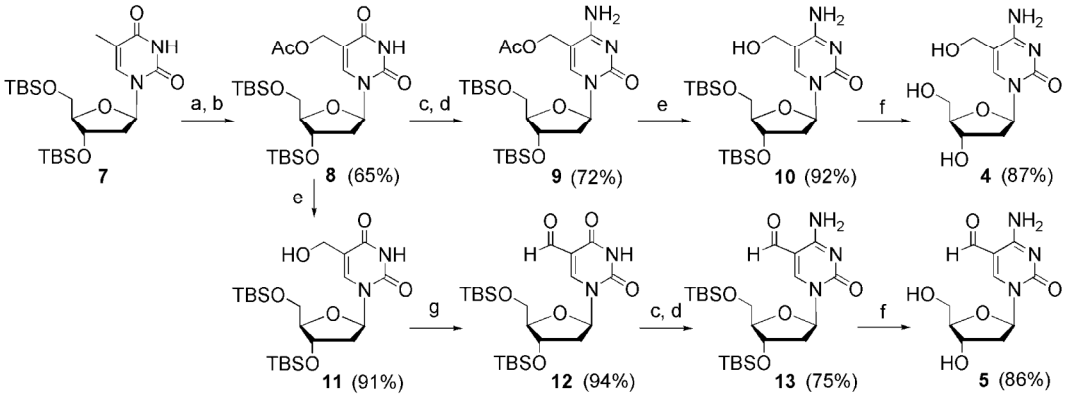
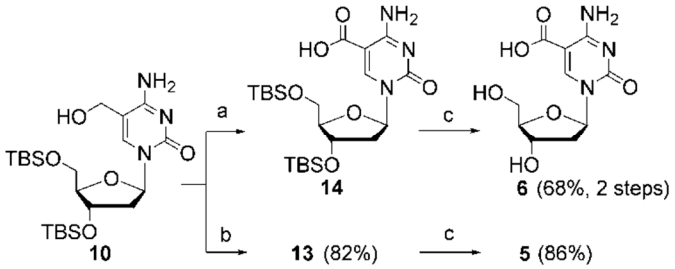


Fig. 1 The $^{5\text{-R}}$dCTP-based PCR technology for the expeditious preparation of long 5-hmC-, 5-fC-, and 5-caC-containing ODNs.[10]



Scheme 1 Synthesis of $^{5\text{-HOMe}}$dC (**4**) and $^{5\text{-CHO}}$dC (**5**). Reagents and conditions: a. NBS, AIBN, CCl_4, 80 °C, 1.5 h; b. CH_3COOK, 40 °C, 30 min; c. TsCl, *N*-methylpiperidine, Et_3N, 4 h; d. 28% NH_4OH, 30 min; e. K_2CO_3/MeOH/H_2O, 20 °C, 2 h; f. THF/H_2O/TFA (2:1:1), 20 °C, 1 h; g. activated MnO_2, CH_2Cl_2, 30 °C, 48 h.

As shown in Scheme 1, 3′,5′-diTBS-protected dT (**7**) which could be easily prepared from dT was brominated with NBS in CCl_4, and then treated with potassium acetate to afford the acetylated intermediate **8** in 65% yield.[12] After **8** was deacetylated with K_2CO_3, mild oxidation of the protected $^{5\text{-HOMe}}$dU (**11**) with activated MnO_2 in CH_2Cl_2 furnished clean conversion to the silylated $^{5\text{-CHO}}$dU (**12**) in 94% yield.[9e,13]



Scheme 2 Synthesis of $^{5\text{-COOH}}$dC (**6**) and $^{5\text{-CHO}}$dC (**5**). Reagents and conditions: a. TEMPO/BAIB (0.2/2.5 equiv), CH_2Cl_2/H_2O, 20 °C, 8 h; b. TEMPO/BAIB (0.2/1.1 equiv), CH_2Cl_2/H_2O, 20 °C, 2 h; c. THF/H_2O/TFA (2:1:1), 20 °C, 1 h.

As Carell mentioned, the major challenge for $^{5\text{-HOMe}}$dCTP (**1**) synthesis is associated with the 5-hydroxylmethyl group in $^{5\text{-HOMe}}$dC (**4**), which makes it difficult to selectively phosphorylate the OH group at C5′ position to access $^{5\text{-HOMe}}$dC 5′-monophosphate (**16**).[10] To address this issue, we attempted to conduct selective amination at C4 position of **8** to obtain the acetylated $^{5\text{-HOMe}}$dC intermediate **9,** which could be later utilized for the synthesis of **16**. However, the conventional $POCl_3$/1,2,4-triazole/ methanolic NH_3 method only gave **9** in low yield (12%) with significant amount of deacetylated product.[4c,14] Switching to 28% NH_4OH largely reduced the undesired deacetylation on the 5-hydroxylmethyl group, and afforded **9** in 60% yield.[12a,15] To further improve the synthetic efficacy and simplify the procedures, TsCl was used with 1-methylpiperidine as the activator to form a more reactive quaternary ammonium intermediate.[16] The in situ aminolysis selectively afforded $^{5\text{-HOMe}}$dC precursor **9** in 72% yield within 30 min. More interestingly, this one-pot reaction with TsCl/1-methylpiperidine/28% NH_4OH exhibited excellent compatibility with the 5-formyl group in **12**, and yielded the protected $^{5\text{-CHO}}$dC precursor **13** in high conversion rate (75%). Final deprotection (deacetylation and desilylation for **9**/desilylation for **13**) gave $^{5\text{-HOMe}}$dC (**4**) and $^{5\text{-CHO}}$dC (**5**) in high yields.

Due to the presence of the TBS groups and glycosidic bond in $^{5\text{-HOMe}}$dC precursor **10**, most conventional strong oxidizing methods are too harsh to transform the hydroxyl group to carboxylic acid with high chemoselectivity. To obtain $^{5\text{-COOH}}$dC (**6**), **10** was oxidized with TEMPO/BAIB (0.2/2.5 equiv) under mild conditions.[17] The subsequent removal of the TBS groups with TFA afforded **6** in nearly 70% yield over two steps (Scheme 2). It is noteworthy that the outcomes of TEMPO/BAIB-mediated oxidation are strongly correlated with the amount of BAIB.[18] When stoichimetric amount of BAIB was applied (TEMPO/BAIB (0.2/1.1 equiv)), **10** could be efficiently transformed into the corresponding aldehyde **13**, providing an alternative approach to $^{5\text{-CHO}}$dC (**5**).

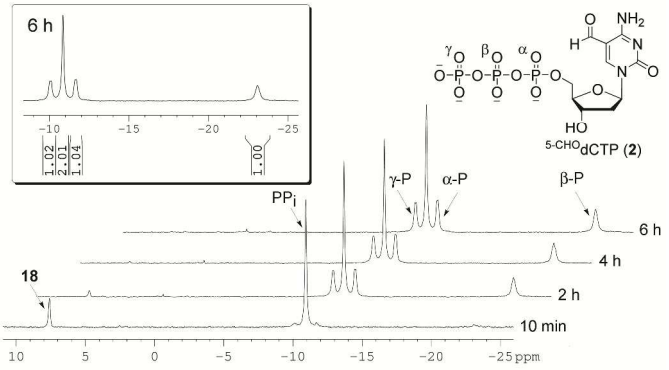
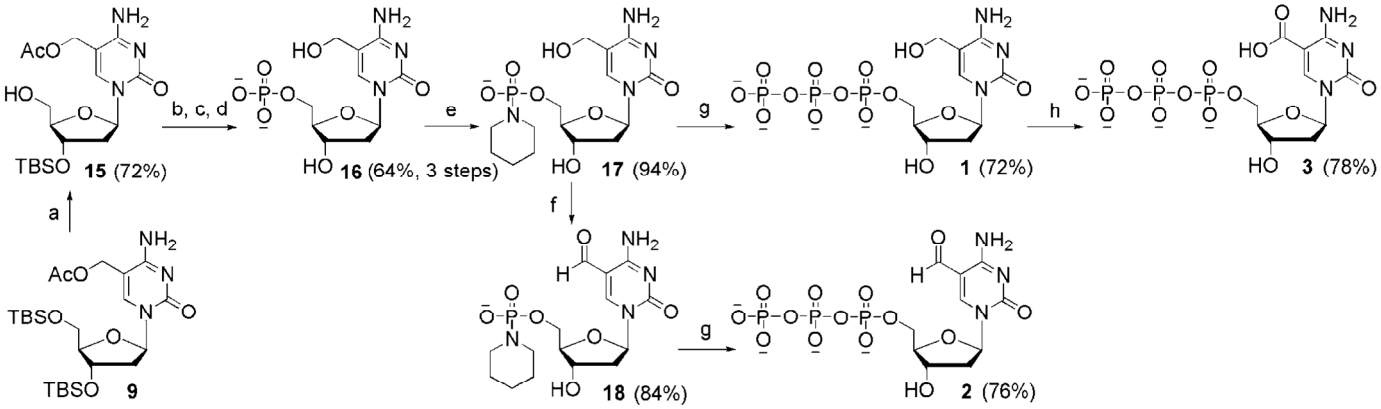


Fig. 2 The stacked ^{31}P NMR tracing spectra of $^{5\text{-CHO}}$dCTP (**2**) synthesis.

Our synthetic route to $^{5\text{-HOMe}}$dCTP (**1**) and $^{5\text{-CHO}}$dCTP (**2**) started from acetylated $^{5\text{-HOMe}}$dC intermediate **9** (Scheme 3). Efficient regioselective desilylation at 5′ position was achieved by lowering the concentration of TFA (THF/H_2O/TFA, 4:1:1) and reaction temperature (0 °C).[19] Treatment of **15** with $POCl_3$ in $PO(OMe)_3$ followed by sequential deacetylation and desilylation afforded the $^{5\text{-HOMe}}$dC 5′-monophosphate (**16**) in 64% yield. Our observation that no proton sponge was required for the phosphorylation of **15** was in agreement with previous reports on the synthesis of cytosine-containing nucleoside



Scheme 3 Synthesis of $^{5\text{-HOMe}}$dCTP (**1**), $^{5\text{-CHO}}$dCTP (**2**), and $^{5\text{-COOH}}$dCTP (**3**). Reagents and conditions: a. THF/H_2O/TFA (4:1:1), 0 °C, 2 h; b. $POCl_3$, $PO(OCH_3)_3$, 0 °C, 2 h; c. TFA/H_2O (1:1), 20 °C, 1 h; d. K_2CO_3/MeOH/H_2O, 20 °C, 1 h; e. 2,2′-dithiodianiline, PPh_3, piperidine, DMSO, 20 °C, 8 h; f. activated MnO_2, MeOH, 50 °C, 24 h; g. $(n\text{Bu}_4\text{N})_3\text{HP}_2\text{O}_7$, DCI, 20 °C, 6 h; h. TEMPO/BAIB (0.4/2.5 equiv), *t*BuOH/CH_2Cl_2/H_2O (4:4:1), 20 °C, 48 h.

monophosphates.[20] In the following step, **16** was converted to $^{5\text{-HOMe}}$dC 5′-phosphoropiperidate (**17**) by the redox condensation method in excellent yield (94%).[11c] Treatment of **17** with activated MnO_2 smoothly oxidized 5-hydroxylmethyl group to 5-formyl group to give $^{5\text{-CHO}}$dC 5′-phosphoropiperidate (**18**). Finally, **17** and **18** were subjected to the 4,5-dicyanoimidazole (DCI)-promoted P(V)–N activation strategy to synthesize $^{5\text{-HOMe}}$dCTP (**1**) and $^{5\text{-CHO}}$dCTP (**2**).[11a,c] ^{31}P NMR tracing experiments showed that both **1** and **2** were obtained with high conversion efficacy as exemplified by the reaction of **2** (Fig. 2), indicating that the P(V)–N activation method well tolerated the hydroxylmethyl and formyl modifications on cytosine. Ethanol precipitation followed by ion exchange chromatography afforded **1** and **2** in high isolated yields.

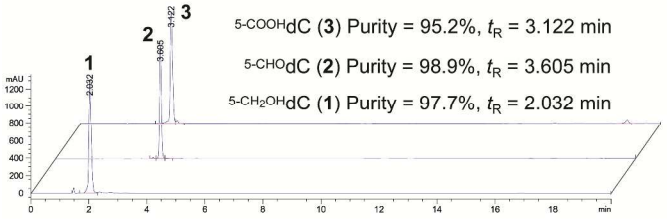


Fig. 3 The HPLC traces of $^{5\text{-HOMe}}$dCTP (**1**), $^{5\text{-CHO}}$dCTP (**2**), and $^{5\text{-COOH}}$dCTP (**3**).

The attempt to oxidize **17** with TEMPO/BAIB (0.2/2.5 equiv) system only afforded the desired $^{5\text{-COOH}}$dC 5′-phosphoropiperidate (**19**) in low yield (<20%) due to the labile nature of phosphoropiperidate under even weakly acidic conditions. Therefore, we directly oxidized $^{5\text{-HOMe}}$dCTP (**1**) with TEMPO/BAIB (0.4/2.5 equiv) to yield $^{5\text{-COOH}}$dCTP (**3**). ^{31}P NMR tracing results showed that the oxidation process in monophasic *t*BuOH/CH_2Cl_2/H_2O (4:4:1) solvent system[21] was smooth and clean. After 48 h, **3** was isolated in 78% yield. The quality of **3** was determined with analytic RP-HPLC along with **1** and **2**. The HPLC traces in Fig. 3 showed that triphosphates **1**–**3** prepared by our method were of high purity (>95%). But it is worth noting that the solvent system also played a key role in the TEMPO/BAIB oxidation. When monophasic CH_3CN/H_2O (1:1) or biphasic CH_2Cl_2/H_2O (1:1) was used, the oxidation of **1** was extremely slow. While the TEMPO/BAIB oxidation of **10** in biphasic CH_2Cl_2/H_2O (1:1) afforded **14** within 8 h, the oxidation in monophasic *t*BuOH/CH_2Cl_2/H_2O (4:4:1) required much longer time (48 h).

Conclusions

In summary, we have developed an efficient method for the preparation of high-quality 5-hydroxylmethyl-, 5-formyl-, and 5-carboxyl-2′-deoxycytidine triphosphates (**1**–**3**) on the basis of the P(V)–N activation strategy. The synthesis of the parent nucleosides (**4**–**6**) were also optimized to provide facile access to all three oxidation products of $^{5\text{-Me}}$dC. The P(V)–N activation method described in this paper along with the $^{5\text{-R}}$dCTP-based PCR technology may greatly facilitate the investigation of $^{5\text{-Me}}$dC-related epigenetic regulations and development of regenerative drugs.

Acknowledgement

We thank the National Natural Science Foundation of China (Nos. 21002041 and 21262014), Key Project of Chinese Ministry of Education (No. 212092), Scientific Research Foundation of Chinese Ministry of Human Resources and Social Security for Returned Chinese Scholars (2011), and Research Funds (Nos. ky2012zy08 and 2013QNBJRC001) and Startup Funds for PhDs (2010) from JXSTNU for financial support.

Notes and references

[a] Jiangxi Key Laboratory of Organic Chemistry, Jiangxi Science & Technology Normal University, 605 Fenglin Avenue, Nanchang, Jiangxi 330013, PR China. Email: sunqi96@tsinghua.org.cn; Tel: +86 791 83805183

[b] Department of Polymer Science & Engineering, College of Chemistry & Chemical Engineering, Nanjing University, Nanjing, Jiangsu 210093, PR China. Email: fengfd@nju.edu.cn; Tel: +86 25 83592119

† Electronic Supplementary Information (ESI) available: Experimental procedures and NMR spectra of intermediates and products are included. See DOI: 10.1039/c000000x/

1 S. Feng, S. E. Jacobsen and W. Reik, *Science*, 2010, **333**, 622.

2 J. A. Law and S. E. Jacobsen, *Nature Rev. Genet.*, 2010, **11**, 204.
3 S. C. Wu and Y. Zhang, *Nature*, 2010, **11**, 607.
4 (a) S. Kriaucionis and N. Heintz, *Science*, 2009, **324**, 929; (b) A. Szwagierczak, S. Bultmann, C. S. Schmidt, F. Spada and H. Leonhardt, *Nucleic Acids Res.*, 2010, **38**, 19; (c) M. Münzel, D. Globisch, T. Brückl, M. Wagner, V. Welzmiller, S. Michalakis, M. Müller, M. Biel and T. Carell, *Angew. Chem. Int. Ed.*, 2010, **49**, 5375; (d) T. Pfaffeneder , B. Hackner, M. Truβ, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeier, T. Carell, *Angew. Chem. Int. Ed.*, 2011, **50**, 7008; (e) C.-X. Song, K. E. Szulwach, Y. Fu, Q. Dai, C.-Q. Yi, X.-K. Li, Y.-J. Li, C.-H. Chen, W. Zhang, X. Jian, J. Wang, L. Zhang, T. J. Looney, B.-C. Zhang, L. A. Godley, L. M. Hicks, B. T. Lahn, P. Jin and C. He, *Nature Biotechnology*, 2011, **29**, 68.
5 (a) M. Tahiliani, K. P. Koh, Y.-H Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind and A. Rao, *Science*, 2009, **324**, 930; (b) T.-P. Gu, F. Guo, H. Yang, H.-P. Wu, G.-F. Xu, W. Liu, Z.-G. Xie, L. Shi, X. He, S.-G. Jin, K. Iqbal, Y.-G. Shi, Z. Deng, P. E. Szabó, G. P. Pfeifer, J. Li and G.-L Xu, *Nature*, 2011, **477**, 606.
6 Y.-F. He, B.-Z. Li, Z. Li, P. Liu, Y. Wang, Q.-Y. Tang, J.-P. Ding, Y.-Y. Jia, Z.-C. Chen, L. Li, Y. Sun, X.-X. Li, Q. Dai, C.-X. Song, K.-L. Zhang, C. He and G.-L. Xu, *Science*, 2011, **333**, 1303.
7 S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers and Y. Zhang, *Nature*, 2010, **466**, 1129.
8 T. P. Jurkowski and A. Jeltsch, *ChemBioChem*, 2011, **12**, 2543.
9 (a) M. Münzel, D. Globisch, C. Trindler and T. Carell, *Org. Lett.*, 2010, **12**, 5671; (b) M. Münzel, U. Lischke, M. D. Stathis, T. Pfaffeneder, F. A. Gnerlich, C. A. Deiml, S. C. Koch, K. Karaghiosoff and T. Carell, *Chem.–Eur. J.*, 2011, **17**, 13782; (c) Q. Dai and C. He, *Org. Lett.*, 2011, **13**, 3446; (d) S. Schiesser, B. Hackner, T. Pfaffeneder, M. Müller, C. Hagemeier, M. Truss and T. Carell, *Angew. Chem. Int. Ed.*, 2012, **51**, 6516; (e) P. Guo, S.-Y. Yan, J.-L Hu, X.-W. Xing, C.-C Wang, X.-W. Xu, X.-Y. Qiu, W. Ma, C.-J Lu, X.-C. Weng and X. Zhou, *Org. Lett.*, 2013, **15**, 3266; (f) A. S. Schrçder, J. Steinbacher, B. Steigenberger, F. A. Gnerlich, S. Schiesser, T. Pfaffeneder and T. Carell, *Angew. Chem. Int. Ed.*, 2014, **53**, 315.
10 B. Steigenberger, S. Schiesser, B. Hackner, C. Brandmayr, S. K. Laube, J. Steinbacher, T. Pfaffeneder and T. Carell, *Org. Lett.*, 2013, **15**, 366.
11 (a) Q. Sun, S.-S. Gong, J. Sun, S. Liu, Q. Xiao and S.-Z. Pu, *J. Org. Chem.* 2013, **78**, 8417; (b) Q. Sun, X.-J. Li, J. Sun, S.-S. Gong, G. Liu and G.-D. Liu, *Tetrahedron*, 2014, **70**, 294; (c) Q. Sun, S.-S. Gong, J. Sun, C.-J. Wang, S. Liu, G.-D. Liu and C. Ma, *Tetrahedron Lett.*, 2014, **55**, 2114; (d) Q. Sun, S.-S. Gong, S. Liu, J. Sun, G.-D. Liu, and C. Ma, *Tetrahedron*, 2014, **70**, 4500.
12 (a) M. de Kort, P. C. de Visser, J. Kurzeck, N. J. Meeuwenoord, G. A. van der Marel, W. Rüger and J. H. van Boom, *Eur. J. Org. Chem.*, 2001, 2075; (b) R. K. Grover, S. J. K. Pond, Q.-Z. Cui, P. Subramaniam, D. A. Case, D. P. Millar and P. W. Jr, *Angew. Chem. Int. Ed.*, 2007, **46**, 2839.
13 C. J. LaFrancois, J. Fujimoto and L. C. Sowers, *Chem. Res. Toxicol.*, 1998, **11**, 75.
14 Adel A.-H. A. Rahman, T. Wada and K. Saigo, *Tetrahedron Lett.*, 2001, **42**, 1061.
15 X.-H. Peng, I. S. Hong, H. Li, M. M. Seidman and M. M. Greenberg, *J. Am. Chem. Soc.*, 2008, **130**, 10299.
16 H. Komatsu, K. Morizane, T. Kohno and H. Tanikawa, *Org. Proc. Res. Dev.*, 2004, **8**, 564.
17 (a) J. B. Epp and T. S. Widlanski, *J. Org. Chem.*, 1999, **64**, 293; (b) L. J. van den Bos, J. D. C. Codée, J. C. van der Toorn, T. J. Boltje, J. H. van Boom, H. S. Overkleeft and G. A. van der Marel, *Org. Lett.*, 2004, **6**, 2165.
18 A. D. Mico, R. Margarita, L. Parlanti, A. Vescovi and G. Piancatelli, *J. Org. Chem.*, 1997, **62**, 6974.
19 X.-F. Zhu, H. J. Williams and A. I. Scott, *J. Chem. Soc., Perkin Trans.1*, 2000, 2305.
20 (a) T. W. Abraham, T. I. Kalman, E. J. McIntee and C. R. Wagner, *J. Med. Chem.*, 1996, **39**, 4569; (b) A. R. Kore, Z.-J. Xiao, A. Senthilvelan, I. Charlesa, M. Shanmugasundaram, S. Mukundarajan and B. Srinivasan, *Nucleos. Nucleot. Nucl.*, 2012, **31**, 567.
21 A. R. Kore, B. Yang and B. Srinivasan, *Tetrahedron Lett.*, 2013, **54**, 5325.

