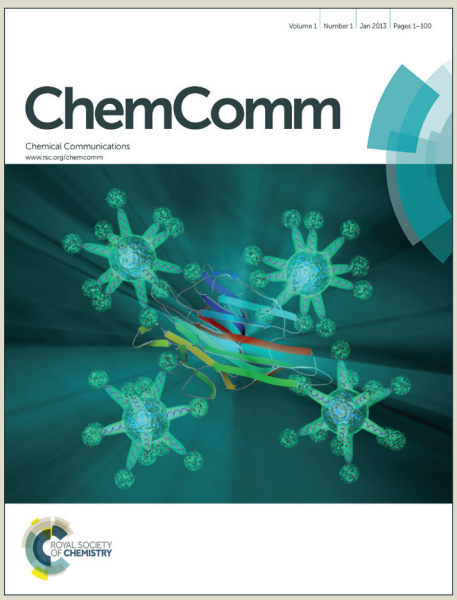
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

RSCPublishing

COMMUNICATION

Water-bridged hydrogen bond formation between 5-hydroxylmethylcytosine (5-hmC) and its 3′-neighbouring bases in A- and B-form DNA Duplexes

Cite this: DOI: 10.1039/x0xx00000x

Received ooth January 2015,
Accepted ooth January 2015

DOI: 10.1039/x0xx00000x

www.rsc.org/

Rui Wang, Srivathsan V. Ranganathan, Vibhav A. Valsankar, Stephanie M. Magliocco, Fusheng Shen, Alan Chen and Jia Sheng*

5-Hydroxylmethylcytosine (5hmC) has been recognized as the sixth base with important biological functions in many tissues and cell types. We present here the high-resolution crystal structures and molecular simulation studies of both A-form and B-form DNA duplexes containing 5hmC. We observed that the 5hmC interacts with its 3′-neighboring bases through water-bridged hydrogen bonds and these interactions may affect the further oxidation of 5hmC.

5-Hydroxylmethylcytosine (5hmC) has been increasingly recognized as the sixth base of the genome besides A, G, C, T and 5mC,[1] although it was firstly identified as a new modified base in some bacterial and viruses in 1950s,[2] and in the animal genome in early 1970s.[3] In 2009, 5hmC was confirmed as another abundant epigenetic modification in embryonic stem cells and Purkinje neurons.[4, 5] Since then, 5hmC has been detected in most of the mammalian tissues and cell types with diverse levels of abundance.[6-10] Although tremendous progresses have been made to study the 5hmC during the past six years,[11-13] its detailed biological functions remain elusive compared to the well-studied epigenetic modification 5mC.[14] To date, 5hmC has been found highly involved in the regulation of gene expression, genome stability and cell development, which are all associated with human diseases like cancer.[15-22] 5hmC is generated through the oxidation of 5-mC by ten-eleven translocation (TET) enzymes,[4] which can also further oxidize the 5hmC to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Figure 1).[18, 23, 24] Other than TET, several other proteins can also recognize and interact with the 5hmC-containing DNA.[25-33]

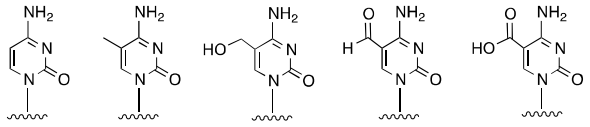


Fig. 1 Chemical structures of DNA nucleosides A, G, T, C, 5mC, 5hmC, 5fC and 5caC.

The recent structural studies have contributed the detailed information about how TET and TET-like enzymes mediate the formation of 5hmC from 5mC through a base-flipping mechanism.[34, 35] In addition, the geometry and base paring properties of 5hmC residue in the B-form Dickerson-Drew dodecamer duplex context have also been studied.[36-39] It was revealed that 5hmC modification does not disrupt the whole B-form double helix, the base pairing geometry, and the overall duplex thermal stability. These findings, in combination with the fact that the regular polymerases can not distinguish the 5-hmC and 5-mC from the natural C, led to the conclusion that the position 5 of cytosine is an ideal place to install epigenetic modifications that are not mutagenic.[37, 40]

More interestingly, the rotation-free hydroxyl group of 5-hmC in these B-form DNA duplex structures uniformly points toward the 3′-end of the strand and interacts with the neighbouring G through water-bridged hydrogen bonds. Considering that this 5-hydroxyl group is located in the major groove of DNA and its orientation will affect the enzyme recognition, it will be interesting to study the stability and dynamic of these hydrogen bonding interactions, as well as their potential effects to the further oxidation of 5hmC. Very recently, the 5hmC residue has also been detected in RNA of mammalian cells and tissues, implying its broader presence and functions in wider range of nucleic acid contexts.[41] Herein, we present two high-resolution crystal structures of both A- and B-form DNA duplexes containing 5hmC. We observed similar interactions between the 5hmC and its 3′-neighboring bases through water-bridged hydrogen bonding in both forms of duplexes. The subsequent molecular simulation work suggests that this type of hydrogen bond is relatively stable. The hydrogen bonding interactions might be related to the bond dissociation energy of the 5-hydroxyl group and the further oxidation step of 5hmC to 5fC.

We chose the previously well studied self-complementary octamer [5′-G(2′-SeMe-dU)GTA(5hmC)AC-3′] as the A-form duplex model to incorporate 5hmC at position 6 with a 5hmC-A step. The 2′-SeMe-dU residue is used to facilitate crystallization and drive this

DNA sequence to A-form duplex without structural perturbation.[42, 43] On the other hand, the typical Dickerson-Drew dodecamer with 5hmC at position 3, [5′-CG(5hmC)GAATTCGCG-3′] with a 5hmC-G step, was used as the ideal B-form duplex model. Considering that the crystallization conditions might affect the molecular packing and the hydration patterns, we crystallized both samples under the same condition (10% MPD, 40 mM Na cacodylate pH 7.0, 12 mM spermine tetra-HCl, 12mM KCl and 80 mM NaCl), which is different as the previously published one.[37] Both DNA crystallized very well within one week and the crystals diffracted to 1.4Å and 1.55Å respectively. The data collection and structure refinement statistics are summarized in Table 1. Both structures were solved by molecular replacement using PDB models of 1Z7I[43]and 1BNA[44] respectively.

Table 1 X-ray data collection and structural refinement statistics of A-form 8mer [5′-dG(2′-SeMe-dU)GTA(5hmC)AC-3′] and B-form Dickerson-Drew dodecamer [5′-dCG(5hmC)GAATTCGCG-3′] containing 5hmC.

DNA (PDB ID)	A-form 8mer (5CH0)	B-form 12mer (5CJY)
Data Collection and Processing		
Space group	$P4_32_12$	$P2_12_12_1$
Unit cell (Å, °)	43.05, 43.05, 23.76 90, 90, 90	25.32, 40.61, 65.40 90, 90, 90
Resolution range, Å (last shell)	50-1.40 (1.42-1.40)	50-1.50 (1.55-1.50)
Unique reflections	4738 (203)	10064 (544)
Completeness, %	99.5 (92.7)	88.6 (50.0)
R_{merge} ,%	6.2 (17.4)	2.6 (27.7)
$\langle I/\sigma(I)\rangle$	55.5 (9.6)	12.3 (4.8)
Redundancy	21.4 (6.6)	14.9 (7.5)
Structure Refinement		
Molecules per asymmetric unit	1 single strand	1 duplex
Resolution range, Å	30.44-1.40	34.5-1.55
Number of reflections	4496	9102
Completeness, %	99.6	92.54
R_{work} , %	18.9	19.9
R_{free} , %	20.1	24.6
Bond length r.m.s. Å	0.01	0.005
Bond angle r.m.s.	1.64	1.16
Overall B-factor, $Å^2$	7.77	21.1

$R_{merge}=\Sigma|I-\langle I\rangle|/\Sigma I$

The overall duplex structures and local 5hmC:G pairing patterns of both DNA are showed in Figure 2. The 5-hydroxylmethyl groups, located in the major groove, turn to the 3′-direction in both structures. Consistent with the previous structure and melting temperature studies,[37, 40, 45] the 5-hmC does not cause obvious duplex structure perturbation in both forms of DNA compared to their native counterparts (Fig. 2A and 2B). When the local base pairing is compared, the 5hmC:G in the B-form 12mer aligns perfectly with the native C:G pair (Fig. 2C). While in the A-form 8mer, the backbone of G paired with 5hmC rotates ~100 degree compared to the native G (Fig. 2D), bringing the overall r.m.s value to 0.5. The direct comparison of two 5hmC residues in these two structures indicates a slight orientation shift (~0.7 Å) of the hydroxyl groups (Fig. 2E), which might be induced by the slightly different hydration patterns and the local interactions between the hydroxyl groups and the connected water molecules.

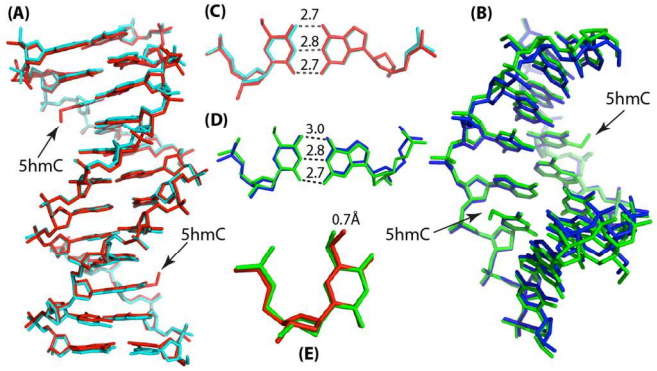


Fig. 2 (A) Duplex superimpose comparison of 5hmC-12mer $[CG(5hmC)GAATTCGCG]_2$ (red) and native-12mer (cyan, PDB: 1BNA) with r.m.s=0.32; (B) Duplex superimpose comparison of 5hmC-8mer [G(2'-SeMe-dU) $GTA(5hmC)AC]_2$ (green) and native-8mer (blue, PDB: 1Z7I) with r.m.s=0.5. (C) Base pair comparison of hmC3-G (red) and native C3-G (cyan). (D) Base pair comparison of hmC6-G (green) and native C6-G (blue). The hydroxylmethyl groups in both DNA are indicated by black arrows. Hydrogen bonds distances are in Å. (E) Superimpose comparison of the two 5hmC residues in 12mer (red) and 8mer (green) DNA duplexes, the distance between the two hydroxyl groups is 0.7Å.

It was reported previously that certain highly conserved water molecules are observed to interact with 5hmC through hydrogen bonding in the Dickerson-Drew dodecamer duplex.[36, 37] These interactions might be a major factor to stabilize the conformation of 5-hydroxylmethyl groups in this duplex. Using different crystallization condition, we also observed the similar water-bridged hydrogen bonding interaction between the hmC3 and G4 in this B-form duplex, as shown in Figure 3A. The bond distance between the O7 of 5hmC and the water W1 is 2.9 Å; and the one between W1 and O6 of G4 is 2.8 Å, indicating these two hydrogen bonds are relatively strong in this form. In addition, W1 also connects to another water molecule W2, which interacts with the O4 atom of the T20 in the complementary sequence with a hydrogen bond length of 2.8 Å. This T20 base pairs with A5 that is connected to G4. This hydration networking might contribute the stability of 5hmC conformation and the overall duplex. Similarly, we also observed the hydrogen bonds between the hmC6 and A7 in the A-form 8mer structure. As shown in Fig. 3B, the highly conserved water molecule W3 has interactions with O7 of 5hmC and N6 of A7 with the hydrogen bond distances of 2.8 Å and 2.9 Å respectively. In addition, this hydration pattern is further expanded to the phosphate oxygen atoms in both 3′- and 5′-end of the 5hmC by two additional water molecules W4 and W5, which might cause slight orientation shift of the 5-hydroxylmethyl group that we observed in Fig. 2E. It is also noteworthy that the hydroxyl groups in both structures form hydrogen bonding with the N4 of 5hmC, which might facilitate the amino-imino tautomerization of N4.[46, 47]

To further check the stability and dynamics of these unique hydrogen bonds connecting 5hmC with its neighbouring bases in different base steps and duplex contexts, we carried out molecular dynamics studies in four structural contexts with 5hmC-A and 5hmC-G step in A- and B-form DNA duplex respectively. The

5hmC-A step in B-form DNA and the 5hmC-G step in A-form DNA were generated by the single base mutation of the current two crystal structures followed by the energy minimization step. We first calculated the orientation of hydroxyl group in 5hmC by monitoring the fluctuations of the dihedral angle ϕ, C5-C6-C7-O7 in the 5hmC. Since the carbon atoms lie in the same plane, the angle indicates the preferential orientation of the hydroxyl group with $\phi < 0$ pointing towards the 5′ side and $\phi > 0$ pointing towards the 3′ side. The equilibrium distribution of ϕ is showed in Figure 4. Interestingly, the hydroxyl group of all the duplexes predominantly points towards the 3′ side as shown by the prominent peak at ~90°. In addition, a significantly smaller peak exists near -120°, showing a weak preference to point towards the 5′ side. A typical orientation dynamics of the hydroxyl group in B-form DNA is showed in Figure S1. The preferential orientation of the hydroxyl group towards the 3′ side stems from a steric clash with the 5′ nucleobase in both A and B forms of the DNA.

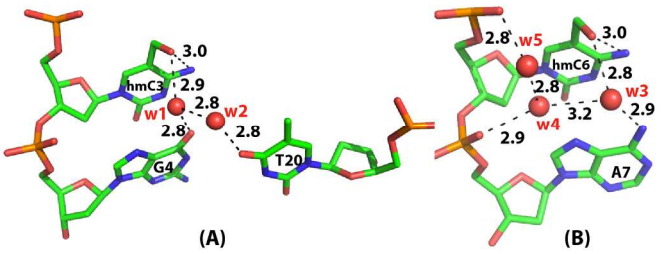


Fig. 3 Stick view of local hydration patterns. (A) 5-hmC3:G4 step and the T20 residue in the complementary strand in B-form 12mer. (B) 5-hmC6:A7 step in A-form 8mer. The red spheres represent the conserved water molecules bridging the hydroxyl group in 5hmC to other atoms of DNA with strong hydrogen bonding.

We also studied the probability distribution of the distances between O7 of 5hmC and N7 of dA or O6 of dG in both A-form and B-form DNA duplexes, as shown in Figure 5. All the four curves show a global peak at the distance of ~4.5 Å, which is ideal for a water-bridged interaction, implying the hydroxyl group orientation is quite stable relative to the 3′-base and the water-bridged hydrogen bonds are most likely maintained throughout the duration of our simulation. More interestingly, there is a significant enhancement in the propensity of the hydroxyl group to directly hydrogen bond with 3′-guanine in the B-form DNA, as evidenced by the peak at ~3 Å. This direct hydrogen-bonding propensity, which is unique in the B-form DNA with a 5hmC-G step, is even more evident if the histogram is normalized by ideal gas reference state using an r-square approximation as shown in the inset (Fig. 5B). This direct hydrogen bond has also been observed in a previous crystal structure where the hydroxyl group in 5hmC is present in a hybrid form.[36] Although in a relatively low abundance, the direct hydrogen bonding might also contribute to the B-form DNA stability. This result also indicates the importance of 3′-nucleobases and DNA geometry to the conformation of 5hmC.

Considering the ubiquity and importance of hydrogen bonding networks in proton-coupled electron transfer of biological systems,[48-50] it is speculated that the water-bridged hydrogen bonds between the 5hmC and the neighbouring bases might affect the pKa of the hydroxyl group, the redox potential of 5hmC and its further oxidation to 5-formylcytidine (5fC), which is another important epigenetic modification.[51] Indeed, in protein oxidation process, it is well known that the hydroxyl group of tyrosine is more inclined to lose the proton when it is involved in the hydrogen bonding interactions with the surrounding residues, which accordingly cause big energy difference.[52, 53] The systematic energy calculation for the oxidation process of 5-hydroxylcytosine with different hydroxyl orientations in varying structural contexts is currently undergoing.

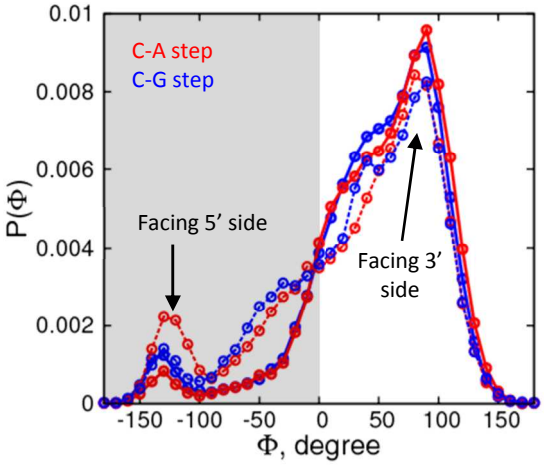


Fig. 4 Dynamic orientation study of the hydroxyl group in C-G and C-A steps in both A-form and B-form DNA duplex by plotting the probability of Φ, the C6-C5-C7-O7 dihedral angle, during the simulation. Positive angle means the OH points to the 3′-side; negative angle means the OH points to the 5′-side. A-form is represented by dashed lines and B-form is represented by smooth lines. C-A step and C-G steps are showed in red and blue respectively.

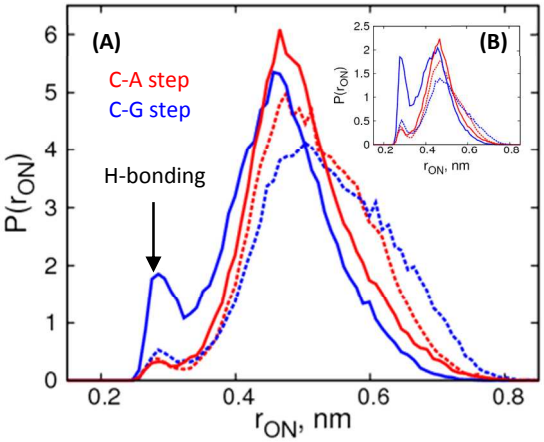


Fig. 5 (A) The probability distribution of the distances (rON) between O7 of 5hmC and N7 of dA or O6 of dG in both A-form and B-form DNA duplex. (B) Same figure with r^2 normalization to roughly approximate ideal gas entropy. A-form is represented by dashed lines and B-form is represented by smooth lines. Higher peak at ~3 Å represents good H-bonding.

In conclusion, through two high-resolution crystal structures and molecular dynamic simulation studies, we observed that in the context of both A-form and B-form DNA duplexes, the 5hmC residue could interact with its 3′-neighbouring bases through relatively stable water-bridged hydrogen bonding. We speculate that these unusual hydrogen bonding interactions might be related to the

bond dissociation energy of the 5-hydroxyl group and the further oxidation step of 5hmC to 5fC.

This work was supported by the start-up fund from SUNY Albany. Rui Wang is currently supported by postdoctoral fellowship from Simons Foundation (338863, R.W.). We are grateful to Prof. Zhen Huang for the discussion and comments, and SeNA Research for the selenium reagents support.

Notes and references

[a] Department of Chemistry, The RNA Institute, University at Albany, State University of New York, 1400 Washington Ave. Albany, NY, 12222. Email: jsheng@albany.edu; Tel: +1 518 437 4419

† Electronic Supplementary Information (ESI) available: Details of experimental and simulation methods are provided. See DOI: 10.1039/c000000x/

1. M. Bachman, S. Uribe-Lewis, X. Yang, M. Williams, A. Murrell and S. Balasubramanian, *Nat. Chem.*, 2014, **6**, 1049.
2. G. R. Wyatt and S. S. Cohen, *Biochem. J.*, 1953, **55**, 774.
3. N. W. Penn, R. Suwalski, C. O'Riley, K. Bojanowski and R. Yura, *Biochem. J.*, 1972, **126**, 781.
4. M. Tahiliani, K. P. Koh, Y. 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.
5. S. Kriaucionis and N. Heintz, *Science*, 2009, **324**, 929.
6. D. Globisch, M. Munzel, M. Muller, S. Michalakis, M. Wagner, S. Koch, T. Bruckl, M. Biel and T. Carell, *PLoS One*, 2010, **5**, e15367.
7. M. Munzel, D. Globisch, T. Bruckl, M. Wagner, V. Welzmiller, S. Michalakis, M. Muller, M. Biel and T. Carell, *Angew. Chem. Int. Ed. Engl.*, 2010, **49**, 5375.
8. C. X. Song, K. E. Szulwach, Y. Fu, Q. Dai, C. Yi, X. Li, Y. Li, C. H. Chen, W. Zhang, X. Jian, J. Wang, L. Zhang, T. J. Looney, B. Zhang, L. A. Godley, L. M. Hicks, B. T. Lahn, P. Jin and C. He, *Nat. Biotechnol.*, 2011, **29**, 68.
9. A. Szwagierczak, S. Bultmann, C. S. Schmidt, F. Spada and H. Leonhardt, *Nucleic Acids Res.*, 2010, **38**, e181.
10. M. Yu, G. C. Hon, K. E. Szulwach, C. X. Song, L. Zhang, A. Kim, X. Li, Q. Dai, Y. Shen, B. Park, J. H. Min, P. Jin, B. Ren and C. He, *Cell*, 2012, **149**, 1368.
11. C. E. Nestor, J. P. Reddington, M. Benson and R. R. Meehan, *Methods Mol. Biol.*, 2014, **1094**, 243.
12. M. Ko, J. An, W. A. Pastor, S. B. Koralov, K. Rajewsky and A. Rao, *Immunol. Rev.*, 2015, **263**, 6.
13. C. X. Song and C. He, *Trends Biochem. Sci.*, 2013, **38**, 480.
14. R. J. Klose and A. P. Bird, *Trends Biochem. Sci.*, 2006, **31**, 89.
15. M. M. Dawlaty, K. Ganz, B. E. Powell, Y. C. Hu, S. Markoulaki, A. W. Cheng, Q. Gao, J. Kim, S. W. Choi, D. C. Page and R. Jaenisch, *Cell Stem Cell*, 2011, **9**, 166.
16. 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. Szabo, G. P. Pfeifer, J. Li and G. L. Xu, *Nature*, 2011, **477**, 606.
17. K. Iqbal, S. G. Jin, G. P. Pfeifer and P. E. Szabo, *Proc. Natl. Acad. Sci. U S A*, 2011, **108**, 3642.
18. S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang, *Science*, 2011, **333**, 1300.
19. M. Ko, Y. Huang, A. M. Jankowska, U. J. Pape, M. Tahiliani, H. S. Bandukwala, J. An, E. D. Lamperti, K. P. Koh, R. Ganetzky, X. S. Liu, L. Aravind, S. Agarwal, J. P. Maciejewski and A. Rao, *Nature*, 2010, **468**, 839.
20. K. P. Koh, A. Yabuuchi, S. Rao, Y. Huang, K. Cunniff, J. Nardone, A. Laiho, M. Tahiliani, C. A. Sommer, G. Mostoslavsky, R. Lahesmaa, S. H. Orkin, S. J. Rodig, G. Q. Daley and A. Rao, *Cell Stem Cell*, 2011, **8**, 200.
21. M. Wossidlo, T. Nakamura, K. Lepikhov, C. J. Marques, V. Zakhartchenko, M. Boiani, J. Arand, T. Nakano, W. Reik and J. Walter, *Nat. Commun.*, 2011, **2**, 241.
22. E. Kriukiene, Z. Liutkeviciute and S. Klimasauskas, *Chem. Soc. Rev.*, 2012, **41**, 6916.
23. Y. F. He, B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He and G. L. Xu, *Science*, 2011, **333**, 1303.
24. T. Pfaffeneder, B. Hackner, M. Truss, M. Munzel, M. Muller, C. A. Deiml, C. Hagemeier and T. Carell, *Angew. Chem. Int. Ed. Engl.*, 2011, **50**, 7008.
25. C. Frauer, T. Hoffmann, S. Bultmann, V. Casa, M. C. Cardoso, I. Antes and H. Leonhardt, *PLoS One*, 2011, **6**, e21306.
26. Z. Liutkeviciute, E. Kriukiene, I. Grigaityte, V. Masevicius and S. Klimasauskas, *Angew. Chem. Int. Ed. Engl.*, 2011, **50**, 2090.
27. O. Yildirim, R. Li, J. H. Hung, P. B. Chen, X. Dong, L. S. Ee, Z. Weng, O. J. Rando and T. G. Fazzio, *Cell*, 2011, **147**, 1498.
28. M. W. Kellinger, C.-X. Song, J. Chong, X.-Y. Lu, C. He and D. Wang, *Nat. Struct. Mol. Biol.*, 2012, **19**, 831.
29. C. S. Nabel, H. Jia, Y. Ye, L. Shen, H. L. Goldschmidt, J. T. Stivers, Y. Zhang and R. M. Kohli, *Nat. Chem. Biol.*, 2012, **8**, 751.
30. M. Mellen, P. Ayata, S. Dewell, S. Kriaucionis and N. Heintz, *Cell*, 2012, **151**, 1417.
31. M. Iurlaro, G. Ficz, D. Oxley, E. A. Raiber, M. Bachman, M. J. Booth, S. Andrews, S. Balasubramanian and W. Reik, *Genome Biol.*, 2013, **14**, R119.
32. C. G. Spruijt, F. Gnerlich, A. H. Smits, T. Pfaffeneder, P. W. Jansen, C. Bauer, M. Munzel, M. Wagner, M. Muller, F. Khan, H. C. Eberl, A. Mensinga, A. B. Brinkman, K. Lephikov, U. Muller, J. Walter, R. Boelens, H. van Ingen, H. Leonhardt, T. Carell and M. Vermeulen, *Cell*, 2013, **152**, 1146.
33. J. R. Horton, J. G. Borgaro, R. M. Griggs, A. Quimby, S. Guan, X. Zhang, G. G. Wilson, Y. Zheng, Z. Zhu and X. Cheng, *Nucleic Acids Res.*, 2014, **42**, 7947.
34. L. Hu, Z. Li, J. Cheng, Q. Rao, W. Gong, M. Liu, Y. G. Shi, J. Zhu, P. Wang and Y. Xu, *Cell*, 2013, **155**, 1545.
35. H. Hashimoto, J. E. Pais, X. Zhang, L. Saleh, Z. Q. Fu, N. Dai, I. R. Correa, Jr., Y. Zheng and X. Cheng, *Nature*, 2014, **506**, 391.
36. L. Lercher, M. A. McDonough, A. H. El-Sagheer, A. Thalhammer, S. Kriaucionis, T. Brown and C. J. Schofield, *Chem. Commun. (Camb)*, 2014, **50**, 1794.
37. D. Renciuk, O. Blacque, M. Vorlickova and B. Spingler, *Nucleic Acids Res.*, 2013, **41**, 9891.
38. M. W. Szulik, P. S. Pallan, B. Nocek, M. Voehler, S. Banerjee, S. Brooks, A. Joachimiak, M. Egli, B. F. Eichman and M. P. Stone, *Biochemistry*, 2015, **54**, 1294.
39. M. Wanunu, D. Cohen-Karni, R. R. Johnson, L. Fields, J. Benner, N. Peterman, Y. Zheng, M. L. Klein and M. Drndic, *J. Am. Chem. Soc.*, 2011, **133**, 486.
40. M. Munzel, U. Lischke, D. Stathis, T. Pfaffeneder, F. A. Gnerlich, C. A. Deiml, S. C. Koch, K. Karaghiosoff and T. Carell, *Chem. Eur. J.*, 2011, **17**, 13782.
41. L. Fu, C. R. Guerrero, N. Zhong, N. J. Amato, Y. Liu, S. Liu, Q. Cai, D. Ji, S. G. Jin, L. J. Niedernhofer, G. P. Pfeifer, G. L. Xu and Y. Wang, *J. Am. Chem. Soc.*, 2014, **136**, 11582.
42. J. Sheng, J. Jiang, J. Salon and Z. Huang, *Org. Lett.*, 2007, **9**, 749.
43. J. Jiang, J. Sheng, N. Carrasco and Z. Huang, *Nucleic Acids Res.*, 2007, **35**, 477.
44. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Proc. Natl. Acad. Sci. U S A*, 1981, **78**, 2179.
45. A. Thalhammer, A. S. Hansen, A. H. El-Sagheer, T. Brown and C. J. Schofield, *Chem. Commun. (Camb)*, 2011, **47**, 5325.
46. N. Karino, Y. Ueno and A. Matsuda, *Nucleic Acids Res.*, 2001, **29**, 2456.
47. H. Hashimoto, S. Hong, A. S. Bhagwat, X. Zhang and X. Cheng, *Nucleic Acids Res.*, 2012, **40**, 10203.
48. D. R. Weinberg, C. J. Gagliardi, J. F. Hull, C. F. Murphy, C. A. Kent, B. C. Westlake, A. Paul, D. H. Ess, D. G. McCafferty and T. J. Meyer, *Chem. Rev.*, 2012, **112**, 4016.
49. S. Y. Reece and D. G. Nocera, *Annu. Rev. Biochem.*, 2009, **78**, 673.
50. E. T. Judd, N. Stein, A. A. Pacheco and S. J. Elliott, *Biochemistry*, 2014, **53**, 5638.
51. M. Bachman, S. Uribe-Lewis, X. Yang, H. E. Burgess, M. Iurlaro, W. Reik, A. Murrell and S. Balasubramanian, *Nat. Chem. Biol.*, 2015, DOI: 10.1038/nchembio.1848.
52. X. Chen, L. Zhang, L. Zhang, J. Wang, H. Liu and Y. Bu, *J. Phys. Chem. B*, 2009, **113**, 16681.
53. V. R. Kaila, M. P. Johansson, D. Sundholm, L. Laakkonen and M. Wistrom, *Biochim. Biophys. Acta.*, 2009, **1787**, 221.

