ChemComm

COMMUNICATION

View Article Online
View Journal

Cite this: DOI: 10.1039/c5cc07354e

Received 2nd September 2015,
Accepted 3rd November 2015

DOI: 10.1039/c5cc07354e

www.rsc.org/chemcomm

The existence of 5-hydroxymethylcytosine and 5-formylcytosine in both DNA and RNA in mammals†

Hao-Ying Zhang,‡ Jun Xiong,‡ Bao-Ling Qi, Yu-Qi Feng* and Bi-Feng Yuan*

We developed a novel strategy by oxidation–derivatization combined mass spectrometry analysis for the determination of 5-hydroxymethylcytosine and 5-formylcytosine in both DNA and RNA. We reported the presence of 5-formylcytosine in RNA of mammals and found that ascorbic acid and hydroquinone can increase the oxidation of 5-methylcytosine to 5-hydroxymethylcytosine in DNA and RNA.

DNA cytosine methylation (5-methylcytosine, 5-mC) is one of the best-characterized epigenetic modifications that plays crucial roles in a variety of cellular processes.[1] Recent studies have demonstrated that active DNA demethylation in mammals can be achieved through a consecutive oxidation of 5-mC by ten–eleven translocation (TET) dioxygenases with the generation of three intermediates, 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-foC), and 5-carboxylcytosine (5-caC).[2–5] In addition to being an intermediate in 5-mC removal, 5-hmC itself is now considered to be an important epigenetic marker that has functions in various physiological processes.[6–10] 5-foC is the further oxidation product of 5-hmC by TET dioxygenases.

Given the rich layers of epigenetic regulation that result from modifications of DNA, reversible RNA modification has been proposed to represent another realm for biological regulation in the form of "RNA epigenetics".[11] Naturally occurring RNA molecules contain various modified nucleosides derived from four standard nucleosides. To date, more than 100 modified nucleosides have been identified in all three kingdoms of life,[12] among which, 5-mC is one of the most important modifications in the RNA molecules with potential functions in the control and regulation of gene transcription and protein translation.[13] Recently, Fu *et al.*[14] reported the existence of 5-hmC for the first time in RNA in mammalian cells and found that TET proteins also possess the activity of catalyzing the formation of 5-hmC from 5-mC in RNA. However, the biological significance of 5-hmC in RNA is less understood.

The 5-hmC content in genomic DNA of mammalian cells can be as low as 0.0004% of cytosines,[15] and 5-foC occurs at a frequency of 1 to 20 per 10^6 cytosines.[16–18] In RNA, the content of 5-hmC can be approximately 1 modification per 10^6 cytosines.[14] Previous bisulfite sequencing data showed that the widespread presence of 5-mC in RNA and 5-mC may dynamically modulate RNA–protein interactions.[13] A recent study suggested that TET enzymes can catalyze the formation of 5-hmC from 5-mC in RNA.[14] Therefore the presence of 5-hmC in RNA and the involvement of TET family enzymes in generating this modification indicate that 5-hmC in RNA may also participate in the epigenetic regulation of gene expression through modulating RNA–protein interactions. However, quantifications of these cytosine modifications are challenging due to their low *in vivo* contents as well as the interference from the highly abundant normal nucleosides. In this respect, sensitive and accurate measurements of these epigenetic modifications are essential to elucidate their functional roles.

In the past decade, considerable advances have been made in the development of methods for the quantification of these cytosine modifications, including radioactive labelling-based thin layer chromatography detection,[3] immunohistochemistry,[19] fluorescence labelling-based detection,[20,21] and liquid chromatography-mass spectrometry (LC-MS).[22,23] Due to its inherent selectivity and sensitivity, LC-MS has been frequently used in the analysis of low abundance modified nucleosides,[16,24] whereas, in these methods, offline HPLC purification was typically required prior to MS analysis to remove the interference from the bulky normal nucleosides, which is laborious and time-consuming. In addition, the ionization efficiencies in electrospray ionization (ESI) of these cytosine modifications are usually low. Thus, it is possible to increase the detection sensitivities during LC-ESI-MS/MS analysis by introducing an easily ionizable moiety into these cytosine modifications.

Key Laboratory of Analytical Chemistry for Biology and Medicine (Ministry of Education), Department of Chemistry, Wuhan University, Wuhan 430072, P. R. China. E-mail: bfyuan@whu.edu.cn, yqfeng@whu.edu.cn; Fax: +86-27-68755595; Tel: +86-27-68755595

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5cc07354e

‡ These authors contributed equally to this work.

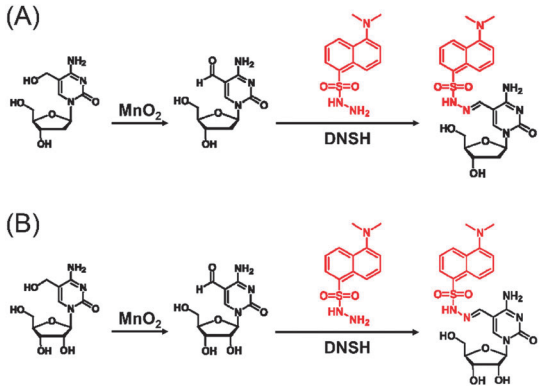

Fig. 1 Oxidation–derivatization reactions of (A) 5-hmdC and (B) 5-hmrC.

In this study, we developed a strategy using oxidation–derivatization combined mass spectrometry (ODMS) analysis for the highly sensitive detection of 5-hydroxymethyl-2′-deoxycytidine (5-hmdC), 5-hydroxymethylcytidine (5-hmrC), 5-formyl-2′-deoxycytidine (5-fodC) and 5-formylcytidine (5-forC). In this strategy, MnO_2 was utilized to oxide 5-hmdC and 5-hmrC to 5-fodC and 5-forC, respectively. The oxidization products were then further labeled by dansylhydrazine (DNSH) with the aldehyde group in 5-fodC and 5-forC (Fig. 1). DNSH harbors the hydrazide moiety that can readily react with aldehyde to give hydrazone derivatives with an easily chargeable moiety (tertiary ammonium), which therefore can increase the ionization efficiency of target analytes during LC-ESI-MS/MS analysis. In addition, the hydrophobic phenyl group introduced by DNSH derivatization can dramatically increase the retention of reversed-phase LC, which, therefore can finally contribute to the increased detection sensitivities.

To evaluate the detection sensitivity of the ODMS method, we performed the experiment under optimized MnO_2 oxidation conditions (Fig. S1, ESI†) and derivatization conditions (Fig. S2, ESI†). The MRM chromatogram shows that the retentions of 5-hmdC and 5-hmrC were relatively weak (~3.2 min) and they coeluted on the C18 reversed-phase column even under optimized separation conditions (Fig. S3, ESI†). However, the retentions of the oxidation–derivatization products of 5-hmdC and 5-hmrC increased dramatically, and the separation resolution also improved (36.4 min for oxidation–derivatization products of 5-hmrC and 38.2 min for oxidation–derivatization products of 5-hmdC) (Fig. S3, ESI†).

The limits of detection (LODs) of the oxidation–derivatization products of 5-hmdC and 5-hmrC were 0.04 and 0.03 fmol, respectively (Table S1, ESI†), which are, to the best of our knowledge, the lowest LODs of 5-hmdC and 5-hmrC ever achieved (the summary of the detection sensitivities of 5-hmdC and 5-hmrC by different methods is shown in Table S1, ESI†). Compared with the native 5-hmdC and 5-hmrC, the detection sensitivities increased by 363-fold and 380-fold for 5-hmdC and 5-hmrC, respectively (Table S2, ESI†).

Using the established ODMS strategy, we quantified 5-hmdC, 5-hmrC, 5-fodC and 5-forC contents in DNA and RNA from HeLa and 293T cells. We first employed high-resolution mass

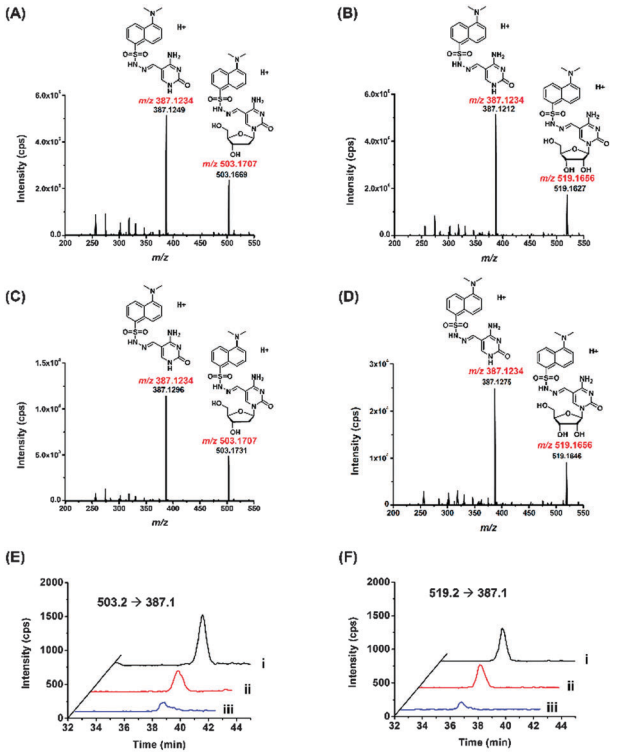

Fig. 2 Determination of endogenous 5-hmdC and 5-hmrC in DNA and RNA. Product ion spectra of oxidation–derivatization product of 5-hmdC standard (A), 5-hmrC standard (B), 5-hmdC detected in HeLa cells (C) and 5-hmrC detected in HeLa cells (D) by high resolution mass spectrometry. The numbers marked in red are the theoretical molecular weights. (E) MRM chromatograms of 5-hmdC standard (i) and 5-hmdC detected in HeLa cells (ii) by ODMS strategy, and endogenous 5-fodC detected in HeLa cells by DNSH derivatization (iii). (F) MRM chromatograms of 5-hmrC standard (i) and 5-hmrC detected in HeLa cells (ii) by ODMS strategy, and endogenous 5-forC detected in HeLa cells by DNSH derivatization (iii).

spectrometry to analyze the oxidation–derivatization products of 5-hmdC and 5-hmrC. The results showed that the product ion spectra of the oxidation–derivatization products of 5-hmdC and 5-hmrC standards (Fig. 2A and B) were identical to that of oxidation–derivatization products of 5-hmdC and 5-hmrC from human HeLa cells (Fig. 2C and D), supporting the existence of 5-hmdC and 5-hmrC in human HeLa cells. And we further explored the lowest amounts of DNA and RNA for the analysis of epigenetic modifications. The results showed that under optimized conditions, 200 ng of DNA and RNA were sufficient for the analysis of endogenous 5-fodC and 5-forC (Fig. S4A and S4B, ESI†). And 20 ng and 100 ng of DNA and RNA were sufficient for the analysis of 5-hmdC and 5-hmrC, respectively (Fig. S4C and S4D, ESI†).

Fig. 2E and F show the typical MRM chromatograms of the oxidation–derivatization products of 5-hmdC and 5-hmrC, respectively. We first quantified the endogenous 5-fodC and 5-forC with DNSH derivatization. The results showed that the measured contents of 5-fodC were $4.5 \pm 0.6/10^6$ dG in HeLa cells and $19.1 \pm 3.3/10^6$ dG in 293T cells (Table S3, ESI†), which are comparable to previously reported results.[16] In addition, we quantified 5-hmdC contents in HeLa cells and 293T cells using

the developed ODMS strategy. In this respect, the endogenous 5-hmdC contents can be obtained by subtracting the endogenous contents of 5-fodC from the total measured contents of 5-fodC. The results showed that the calculated contents of 5-hmdC in HeLa and 293T cells were 68.4 ± 10.1/10^6 dG and 109.1 ± 14.4/10^6 dG, respectively (Table S3, ESI†), which are consistent with previously measured contents.[15]

Similarly, the measured contents of 5-hmrC in RNA were 3.1 ± 0.4/10^6 rG in HeLa cells and 6.9 ± 0.7/10^6 rG in 293T cells (Table S2, ESI†), which are comparable with previously reported results.[14] And the measured contents of 5-forC were 9.0 ± 1.2/10^6 rG in HeLa cells and 8.5 ± 1.4/10^6 rG in 293T cells (Table S3, ESI†). It is worth noting that 5-forC, for the first time, reported the presence in RNA in mammalian cells. And the biological functions of 5-forC in RNA need further investigation.

Ascorbic acid is one of the most important nutrients for mammals. Recently, it has been demonstrated that ascorbic acid could enhance the catalytic activity of TET dioxygenases for the oxidation of 5-mdC to 5-hmdC.[25,26] In addition, another study revealed that hydroquinone can also promote the generation of 5-hmdC in a mechanism involving the elevated activity of TET dioxygenases.[27] Considering the similarity of the chemical structure between 5-hmdC and 5-hmrC, we speculate that ascorbic acid and hydroquinone may also increase the content of 5-hmrC in RNA since the previous study demonstrated that TET dioxygenases could oxidize 5-mrC to 5-hmrC in a similar way to TET dioxygenases converting 5-mdC to 5-hmdC.[14] In this respect, here we used ascorbic acid and hydroquinone treated HeLa and 293T cells and simultaneously examined the change in the contents of 5-hmdC and 5-hmrC in DNA and RNA.

As shown in Fig. 3A, both ascorbic acid and hydroquinone treatment induced a significant increase of 5-hmdC in HeLa cells (2.4-fold increase for ascorbic acid treatment, $p = 0.002$; 1.7-fold increase for hydroquinone treatment, $p = 0.009$) and 293T cells (2.7-fold increase for ascorbic acid treatment, $p = 6.5 \times 10^{-4}$; 1.3-fold increase for hydroquinone treatment, $p = 0.017$), which is consistent with previous studies.[25,27] Similarly, 5-hmrC in RNA also significantly increased in HeLa cells (4.5-fold increase for ascorbic acid treatment, $p = 2.7 \times 10^{-4}$; 2.3-fold increase for hydroquinone treatment, $p = 4.0 \times 10^{-4}$) and 293T cells (3.3-fold increase for ascorbic acid treatment, $p = 0.002$; 1.6-fold increase for hydroquinone treatment, $p = 0.022$) by ascorbic acid and hydroquinone treatment (Fig. 3B). The results suggested that the induced increase of 5-hmrC in RNA by ascorbic acid and hydroquinone treatments might go through the same mechanism as that of 5-hmdC, *i.e.*, the elevated catalytic activity of TET dioxygenases caused by ascorbic acid and hydroquinone can enhance both 5-hmdC and 5-hmrC contents in DNA and RNA.

As for the content of 5-fodC upon ascorbic acid and hydroquinone treatments, significant increases were observed in HeLa cells (3.1-fold increase for ascorbic acid treatment, $p = 0.005$; 4.1-fold increase for hydroquinone treatment, $p = 3.5 \times 10^{-5}$) and 293T cells (2.9-fold increase for ascorbic acid treatment, $p = 0.004$; 2.2-fold increase for hydroquinone treatment, $p = 0.007$) (Fig. 3C). Similarly, significant increases were also observed for 5-forC in HeLa cells (2.7-fold increase for ascorbic acid treatment, $p = 0.003$; 2.0-fold increase for hydroquinone treatment, $p = 0.012$) and 293T cells (2.6-fold increase for ascorbic acid treatment, $p = 0.001$; 2.3-fold increase for hydroquinone treatment, $p = 1.9 \times 10^{-4}$) (Fig. 3D). Since 5-fodC is produced from 5-hmdC by TET dioxygenases, both the increased activity of TET dioxygenases and the elevated 5-hmdC contents could contribute to the increased contents of 5-fodC upon ascorbic acid and hydroquinone treatments. And similar phenomena were observed for 5-forC, further indicating that the endogenous 5-forC may be generated through the oxidation of 5-hmrC by TET dioxygenases.

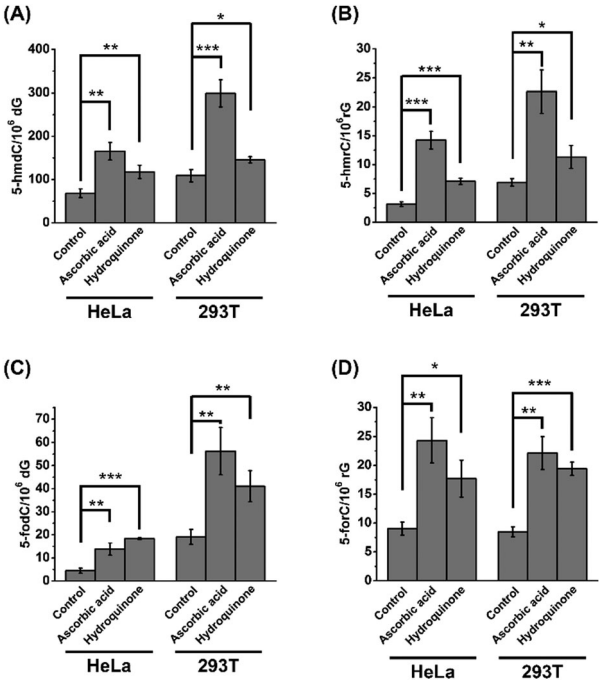

Fig. 3 Contents change of (A) 5-hmdC, (B) 5-hmrC, (C) 5-fodC and (D) 5-forC in HeLa and 293T cells by ascorbic acid and hydroquinone treatments. Ascorbic acid and hydroquinone concentrations, 100 μM; treatment time, 24 h. Error bars represent the SEM ($n = 6$); *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

The discovery of 5-forC in cellular RNA in the current study together with the recently identified presence of 5-hmrC in RNA suggest that the function of TET dioxygenases are not limited to the epigenetic regulation at the DNA level, but can also be extended to RNA. The present work sets the stage for future studies in defining the site-specific localization of 5-forC in RNA as well as the specific binding proteins, and the functions of 5-mrC oxidation in RNA biology. To fully understand the mechanism of these modifications on gene expression regulations, it is essential to develop a high resolution sequencing method to map the distribution of these modifications.

In the current study, we developed a highly sensitive method to simultaneously determine 5-hmC and 5-foC in both DNA and RNA by MnO_2 oxidation and DNSH derivatization coupled with mass spectrometry (ODMS) analysis. Upon MnO_2 oxidation, 5-hmdC and 5-hmrC were converted to 5-fodC and 5-forC, respectively, which were then further labeled by DNSH. DNSH derivatization notably improved the retentions of these cytosine

modifications on reversed-phase chromatography and dramatically increased the detection sensitivities of 5-hmdC and 5-hmrC by 363- and 380-fold, respectively. Using this ODMS method, we successfully quantified 5-hmdC, 5-hmrC, 5-fodC and 5-forC in DNA and RNA of mammalian cells. We reported for the first time the presence of endogenous 5-forC in RNA of mammalian cells. In addition, we found that both ascorbic acid and hydroquinone treatment of the mammalian cells can stimulate the oxidation of 5-mC to 5-hmC in both DNA and RNA. The developed ODMS strategy is an efficient method for the sensitive and simultaneous determination of 5-hmC and 5-foC in both DNA and RNA, which can promote the functional studies of epigenetic modifications.

The authors thank the financial support from the National Basic Research Program of China (973 Program) (2013CB910702, 2012CB720601), and the National Natural Science Foundation of China (21522507, 21475098).

References

1 G. Egger, G. Liang, A. Aparicio and P. A. Jones, *Nature*, 2004, **429**, 457–463.
2 L. Shen, C. X. Song, C. He and Y. Zhang, *Annu. Rev. Biochem.*, 2014, **83**, 585–614.
3 S. Kriaucionis and N. Heintz, *Science*, 2009, **324**, 929–930.
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**, 930–935.
5 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**, 1303–1307.
6 M. R. Branco, G. Ficz and W. Reik, *Nat. Rev. Genet.*, 2012, **13**, 7–13.
7 Y. Tang, S. J. Zheng, C. B. Qi, Y. Q. Feng and B. F. Yuan, *Anal. Chem.*, 2015, **87**, 3445–3452.
8 M. L. Chen, F. Shen, W. Huang, J. H. Qi, Y. Wang, Y. Q. Feng, S. M. Liu and B. F. Yuan, *Clin. Chem.*, 2013, **59**, 824–832.
9 S. G. Jin, Y. Jiang, R. Qiu, T. A. Rauch, Y. Wang, G. Schackert, D. Krex, Q. Lu and G. P. Pfeifer, *Cancer Res.*, 2011, **71**, 7360–7365.
10 J. Xiong, H. P. Jiang, C. Y. Peng, Q. Y. Deng, M. D. Lan, H. Zeng, F. Zheng, Y. Q. Feng and B. F. Yuan, *Clin. Epigenet.*, 2015, **7**, 72.
11 C. He, *Nat. Chem. Biol.*, 2010, **6**, 863–865.
12 M. A. Machnicka, K. Milanowska, O. Osman Oglou, E. Purta, M. Kurkowska, A. Olchowik, W. Januszewski, S. Kalinowski, S. Dunin-Horkawicz, K. M. Rother, M. Helm, J. M. Bujnicki and H. Grosjean, *Nucleic Acids Res.*, 2013, **41**, D262–D267.
13 J. E. Squires, H. R. Patel, M. Nousch, T. Sibbritt, D. T. Humphreys, B. J. Parker, C. M. Suter and T. Preiss, *Nucleic Acids Res.*, 2012, **40**, 5023–5033.
14 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**, 11582–11585.
15 Y. Tang, J. M. Chu, W. Huang, J. Xiong, X. W. Xing, X. Zhou, Y. Q. Feng and B. F. Yuan, *Anal. Chem.*, 2013, **85**, 6129–6135.
16 S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang, *Science*, 2011, **333**, 1300–1303.
17 B. F. Yuan and Y. Q. Feng, *TrAC, Trends Anal. Chem.*, 2014, **54**, 24–35.
18 B. F. Yuan, *Adv. Clin. Chem.*, 2014, **67**, 151–187.
19 M. C. Haffner, A. Chaux, A. K. Meeker, D. M. Esopi, J. Gerber, L. G. Pellakuru, A. Toubaji, P. Argani, C. Iacobuzio-Donahue, W. G. Nelson, G. J. Netto, A. M. De Marzo and S. Yegnasubramanian, *Oncotarget*, 2011, **2**, 627–637.
20 J. Hu, X. Xing, X. Xu, F. Wu, P. Guo, S. Yan, Z. Xu, J. Xu, X. Weng and X. Zhou, *Chem. – Eur. J.*, 2013, **19**, 5836–5840.
21 P. Guo, S. Yan, J. Hu, X. Xing, C. Wang, X. Xu, X. Qiu, W. Ma, C. Lu, X. Weng and X. Zhou, *Org. Lett.*, 2013, **15**, 3266–3269.
22 Y. Tang, J. Xiong, H. P. Jiang, S. J. Zheng, Y. Q. Feng and B. F. Yuan, *Anal. Chem.*, 2014, **86**, 7764–7772.
23 R. Yin, J. Mo, M. Lu and H. Wang, *Anal. Chem.*, 2015, **87**, 1846–1852.
24 S. Liu, J. Wang, Y. Su, C. Guerrero, Y. Zeng, D. Mitra, P. J. Brooks, D. E. Fisher, H. Song and Y. Wang, *Nucleic Acids Res.*, 2013, **41**, 6421–6429.
25 R. Yin, S. Q. Mao, B. Zhao, Z. Chong, Y. Yang, C. Zhao, D. Zhang, H. Huang, J. Gao, Z. Li, Y. Jiao, C. Li, S. Liu, D. Wu, W. Gu, Y. G. Yang, G. L. Xu and H. Wang, *J. Am. Chem. Soc.*, 2013, **135**, 10396–10403.
26 K. Blaschke, K. T. Ebata, M. M. Karimi, J. A. Zepeda-Martinez, P. Goyal, S. Mahapatra, A. Tam, D. J. Laird, M. Hirst, A. Rao, M. C. Lorincz and M. Ramalho-Santos, *Nature*, 2013, **500**, 222–226.
27 J. B. Coulter, C. M. O'Driscoll and J. P. Bressler, *J. Biol. Chem.*, 2013, **288**, 28792–28800.

