RSC Advances

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2022, **12**, 19885

Rapid and highly specific detection of site-specific 5-hydroxymethylcytosine based on peroxotungstate oxidation and mismatch ligation-based LAMP†

Zhenhao Zhang,[ab] Tong He,[ab] Yan Qi,[a] Yuxuan Dai,[a] Kejing Lao*[a] and Xingchun Gou *[a]

Received 26th May 2022
Accepted 21st June 2022

DOI: 10.1039/d2ra03310k

rsc.li/rsc-advances

We have developed a rapid and specific method for site-specific 5-hydroxymethylcytosine (5hmC) quantification at single-base resolution. This bisulfite-free method integrates the peroxotungstate oxidization with the mismatched probe-assisted ligation to guarantee the specificity. Moreover, the high-efficiency LAMP also makes the proposed method suitable for the detection of low-content samples.

DNA methylation plays a pivotal role in regulating gene expression, and is closely related to numerous physiological and pathological processes.[1–3] 5hmC is an oxidation derivative of 5-methylcystein (5mC) catalysed by ten-eleven translocation (TET) family proteins and can be further oxidized to 5-formylcytosine and 5-carboxylcytosine during demethylation.[4,5] The majority of 5hmC are stable DNA modification and possess tissue specificity, which are enriched in embryonic stem cells and brain tissue.[6–9] 5hmC not only plays an important role in neurodevelopment and disease but also involves in the regulation of chromatin structure and gene expression.[10] Notably, the level of 5hmC is markedly depleted in tumors compared with normal tissues.[11] Therefore, 5hmC has been considered as an epigenetic biomarker.[12]

In view of the significance of 5hmC in the epigenetic regulation and the early diagnosis of diseases, a collection of genome-wide 5hm detection strategies have been well developed. For example, high-performance liquid chromatography (HPLC) and mass spectrometry (MS) are widely used in determining the genomic-wide content of 5hmC in all cytosine residues.[13,14] Nevertheless, sequence information and the position of 5hmC could not be obtained. In comparison, the position of 5hmC in genomic DNA can be approximately mapped by affinity enrichment or chemical conversion-assisted sequencing strategies.[15–20] Specifically, oxidative bisulfite sequencing (oxBS-seq) has shown advantage of discriminating 5hmC at single-base resolution.[21] Based on oxidative bisulfite conversion, HpaII-mediated ligation PCR is reported which can detect site-specific 5hmC with high sensitivity and specificity.[22] But they involve bisulfite treatment, which can lead to DNA sequence degradation and require high-input sample. Therefore, it is highly desirable to develop a bisulfite-free, rapid and specific method for accurate quantification of site-specific 5hmC in routine laboratory and clinical applications. The single-molecule immunofluorescent imaging can sensitively detect 5hmC in genomic DNA and provide rapid readout.[23] Moreover, the peroxotungstate oxidation-mediated two-phase amplification systems is bisulfite-free and could facilitate the acquisition of 5hmC level.[24] Whereas, they require multiple purification procedures to reduce background interference, thus leading to a large loss of DNA information.

Generally, there are three obstacles to rapid, simple, and high specific quantification of 5hmC at specific sites. They are interference of C and 5mC, DNA degradation, and loss of DNA information, respectively. Consequently, we develop a label-free and bisulfite-free method for the quantification of site-specific 5hmC at single-base resolution in low-input samples. Due to the advantages of high efficiency and mildness, peroxotungstate is used as a conversion reagent that can selectively oxidate 5hmC to trihydroxylatedthymine (thT) in one step while leaving C and 5mC unchanged.[25–27] To selectively distinguish thT from C and 5mC, ligation-based loop-mediated isothermal amplification (LAMP) is applied,[28,29] since it is prior to primer extension reaction with respect to discern differ of single base.[30,31] Nevertheless, in ligation-based LAMP, it is difficult to completely eliminate non-specific ligation by relying only on ligase to recognize the difference of a single base.[32–34] In this consideration, we designed a mismatched ligation strategy in which a mismatched base with target sequence is introduced at the third position from the 3′-end of probe B. The mismatched

[a]*Shaanxi Key Laboratory of Brain Disorders, Shaanxi Provincial Key Laboratory of Infection and Immunity Diseases, Institute of Basic and Translational Medicine, Xi'an Medical University, No. 1 Xin Wang Road, Xi'an 710021, China. E-mail: gouxingchun@189.cn; Tel: +86 29 86177603*

[b]*School of Chemistry and Chemical Engineering, Shaanxi Normal University, Xi'an 710062, Shaanxi Province, P. R. China*

† Electronic supplementary information (ESI) available. See https://doi.org/10.1039/d2ra03310k

base cooperates with ligase to endow the mismatch ligation strategy with high specificity. There are several distinct merits of the proposed method. In the first place, the one-step oxidation greatly saves time for sample treatment and reduces the DNA loss of the sample. Secondly, mismatch ligation-based LAMP is characterized by high sensitivity and specificity. It is extremely suitable for clinical detection. Furthermore, the performance of this method in analysing low-input samples is highlighted by comparing with ligation-based LAMP and gap ligation-based LAMP.

The principle of the mismatch ligation-based LAMP assay for detection site-specific 5hmC is shown in Fig. 1. We synthesized three double-stranded DNA (dsDNA) target sequences and the detection site is C, 5mC and 5hmC, respectively. Therefore, the three targets are named target C, target 5mC and target 5hmC, respectively. The targets are oxidized with peroxotungstate, and target 5hmC is oxidized to form thT which is prefer to pair with adenine (A), but C and 5mC unchanged. So that target 5hmC can be distinguished from target C, and 5mC by ligation reaction. In the ligating reaction, probe A and probe B are designed to specifically recognize the target thT sequence and then initiate ligation reaction. The high specificity probe B is introduced to the mismatch ligation platform. Base A is designed at the 3′-end of probe B to distinguish the converted C, 5mC and 5hmC. A mismatched base guanine (G) is introduced at the third position from the 3′-end of probe B which cooperates with the terminal base to prevent the ligation of C, and 5mC. Ultimately, probe B can form a pair of mismatched bases (G–G) with the target thT, and two pairs of mismatched bases (A–C, G–G and A–5mC, G–G) with target C and 5mC, respectively. The green part of probe A is complementary to the target sequence and modified with a phosphate group at its 5′-terminus. The dark blue stem loop structures of probe A and probe B are the universal sequences which can be used to initiate the LAMP amplification with a pair of universal primers (shown in Fig. S1†). The probe A and probe B can bind to the target sequence and be adjacent to each other. Only target thT is capable of effectively driving the ligation of probes A and B by Ampligase to form a double stem-loop structure, which is the initiate substance of LAMP. During the ligation process, the mismatched probe B can effectively reduce non-specific ligation and improve the specificity of the method.

The analytical performance of the mismatch ligation-based LAMP strategy is evaluated by tracking the fluorescence in response to different concentrations of target 5hmC. As shown in Fig. 2, the point of inflection (POI) value is the times corresponding to the maximum slope in real-time fluorescence curve, which is linearly dependent on the logarithm (lg) of concentration of target 5hmC from 2 fM to 200 pM. The correlation equation is POI = -5.83 lg C_{5hmC}-27.7 and the corresponding correlation coefficient R is 0.999. To evaluate the specificity of the mismatch ligation-based LAMP strategy, 200 fM of targets C, 5mC, and 5hmC are tested in the same conditions. According to the correlation equation (Fig. 2b), the interference signals produced by target C and target 5mC are estimated to be 0.8, and 1.2% compared with 200 fM target 5hmC, respectively. Compared with the ligation-based LAMP (Fig. S2†), this method improved specificity by 3.7% and 8.6% in the detection of target C and 5mC, respectively.

To further verify the advantage of mismatch ligation-based LAMP in detecting site-specific 5hmC in low-input sample, we designed gap ligation-based LAMP. In gap ligation-based LAMP (Fig. 3), the probe A is same as mismatch ligation-based LAMP. However, probe B is loss of A at the 3′-end, so it is named probe B2. Probes A and B2 are hybridized to target flanking the detection site and left a single-base gap opposite the detection site. When dATP is added to the reaction solution, jumpStart™ Taq DNA polymerase selectively catalyze the base A extension at 3′ end of probe B2 relying on target thT. The base-incorporated probe B2 could then be covalently joined with probe A by Ampligase. Subsequently, the amplification reaction is advanced. But probe B2 is unable to perform a single-base

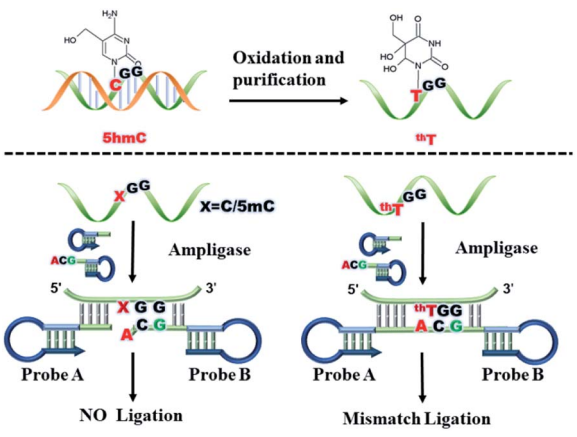

Fig. 1 Schematic representation of the mismatch ligation-based LAMP strategy.

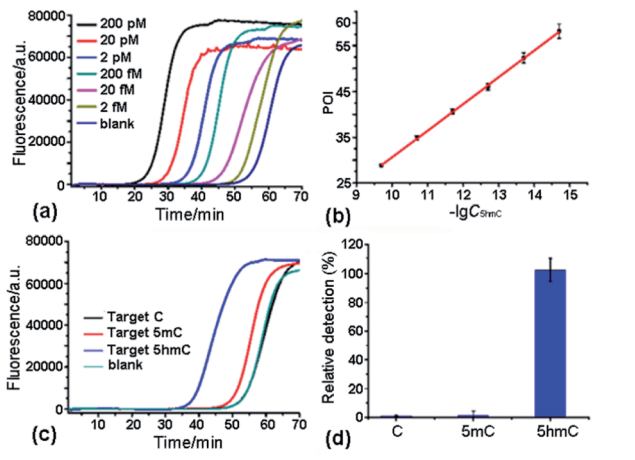

Fig. 2 Dynamic range and sensitivity of LAMP based on mismatched ligation assay. (a) Real time amplification curve produced by synthetic target 5hmC from right to left, the concentration of target 5hmC successively is blank, 2 fM, 20 fM, 200 fM, 2 pM, 20 pM, 200 pM. (b) The relationship between POI and lg of target 5hmC concentration (M). (c) Real time amplification curve produced by 20 fM synthetic target C, target 5mC, and target 5hmC. (d) Relative responses of C and 5mC compared with that of 5hmC.

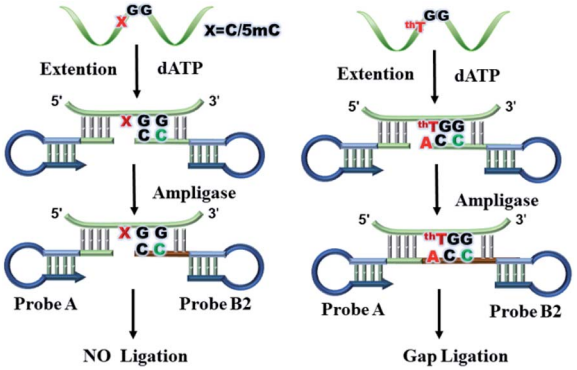

Fig. 3 Schematic representation of gap ligation-based LAMP assay.

extension at the 3′ end depending on target C and 5mC due to the mismatch of targets with the extension base.

We investigate the sensitivity and specificity of the gap ligation-based LAMP. As shown in Fig. 4, the POI values are linearly dependent on the lg of concentrations of target 5hmC in the range from 20 fM to 200 pM. The correlation equation is POI = $-6.47 \lg C_{5hmC}$-33.5 and the corresponding correlation coefficient R is 0.985. To evaluate the specificity of the gap ligation-based LAMP assay, targets C, 5mC and 5hmC at 2 pM are tested under the same conditions. According to the correlation equation (Fig. 4b), the signal interference generated by target C and target 5mC is estimated to be 1.4% and 6.9%, respectively. Compared with mismatch ligation-based LAMP, the specificity and sensitivity of gap ligation-based LAMP are not improved. This may be due to the lower efficiency of probe extension and ligation at low levels of target DNA.

Considering the sensitivity and specificity, mismatch ligation-based LAMP is more suitable for detecting low-content 5hmC. In order to prove the usability of method, the method is

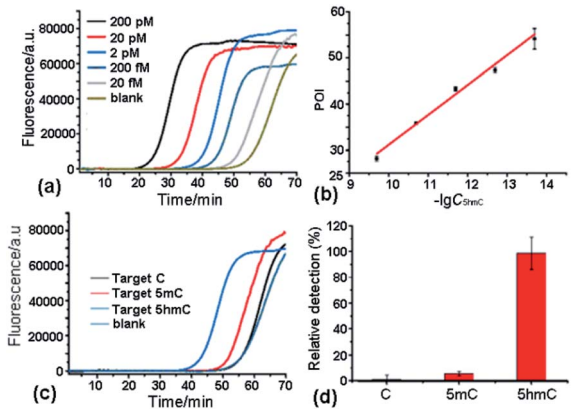

Fig. 4 The dynamic range and sensitivity of LAMP based on gap-ligation assay. (a) Real time amplification curve produced by synthetic target 5hmC from right to left, the concentration of target 5hmC successively is blank, 20 fM, 200 fM, 2 pM, 20 pM, 200 pM. (b) The relationship between POI and lg of target 5hmC concentration (M). (c) Real time amplification curve produced by 20 fM synthetic target C, target 5mC and target 5hmC. (d) Relative responses of C, and 5mC compared with that of 5hmC.

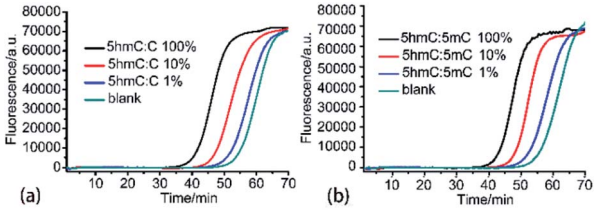

Fig. 5 (a) Real time amplification curve produced by of target 5hmC at different ratio in the mixture of target 5hmC, and target C with the concentration of 200 fM. (b) Real time amplification curve produced by target 5hmC at different ratio in the mixture of target 5hmC, and target 5mC with the concentration of 100 fM.

applied to detect 5hmC in the mixture which contained a large excess of target C or target 5mC. Target 5hmC and target C are mixed for a total concentration of 200 fM. Target 5hmC and target 5mC are mixed for a total concentration of 100 fM as the tested samples, in which the ratio of target 5hmC varied from 1% to 100%. The results are shown in Fig. 5. One can see from Fig. 5a, the target 5hmC can be detected as low as 1% with a real-time fluorescent signal in mixtures containing 99-fold higher target C. Accordingly, the POI value of the fluorescence curve of target 5hmC is calculated by the correlation equation of Fig. 2b. They are 200 fM, 20 fM, 2 fM, and SD are 1.5%, 2.8%, and 3.34%, respectively, indicating that the target C does not interfere with the detection of 5hmC. Detection of mixtures containing target 5mC and 5hmC as described above. The POI of target 5hmC is calculated by the correlation equation of Fig. 2b. They are 100 fM, 10 fM, 1 fM, and SD are 1.52%, 1.2%, and 2.5%, respectively, indicating that the target 5mC does not interfere with the detection of 5hmC. Therefore, detection of small amount of target 5hmC in the presence of large excess of target C or target 5mC is feasible with the proposed assay.

Further studies have been performed to test the practical capability of this assay in genomic DNA. As shown in Fig. 6, we have detected the amount of 5hmC which is located in low density lipoprotein receptor-related protein gene (Lrpab1) in mouse brain genomic DNA. It can be concluded from the correlation equation in Fig. 2b that 10 fM of site-specific 5hmC is contained in 300 ng of genomic DNA. The standard deviation

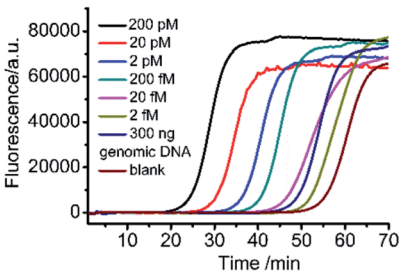

Fig. 6 Analysis of 5hmC in mouse brain genomic DNA. Real time amplification curve produced by target 5hmC and 300 ng mouse brain genomic DNA. Except blank, from left to right, the concentrate on of target 5hmC successively is 200 pM, 20 pM, 2 pM, 200 fM, 20 fM, 300 ng mouse brain genomic DNA, 2 fM, blank.

(SD) and coefficient of variation (RSD) are 1.52, and 3.03%, respectively. To further confirm the results, 300 ng of mouse genomic DNA is spiked with 50 fM synthetic 5hmC and measured in three replicates, resulting in an average recovery of 109%. Therefore, the proposed assay can accurately quantify 5hmC down to fM levels in genomic DNA samples.

Conclusions

In summary, we have developed a highly specific method for the rapid detection of site-specific 5hmc at single-base resolution by combining the peroxotungstate oxidation and mismatched ligation-LAMP. In this design, 5hmC can be converted to thT by only one-step oxidation, and has a single-base difference from C and 5mC. Accordingly, a probe (probe B) with a mismatched base relative to the target sequence at the third base of 3′ end is designed to greatly reduce the generation of non-specific ligation, allowing discriminate as low as 1% target 5hmC in the mixture. In this method, DNA degradation and DNA loss are reduced as bisulfite and multiple purification treatments are not required. Hence, this method can detect target 5hmC down to 2 fM with efficient DNA amplification and obtain detection results within 5 hours, which is of great significance for 5hmC-related clinical studies. Furthermore, this method is applicable to different sequences by changing the complementary sequence of the probe to the target sequence, which avoids the delicate design of the LAMP primers, and significantly improves practicability. In conclusion, the one-step oxidation mediated mismatch ligation-based LAMP method may act as a powerful tool to speed up our understanding of the occurrence of disease at the molecular level.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors are grateful to the Scientific Research Program funded by Shaanxi Provincial Education Department (21JS037), Natural Science Basic Research Plan in Shaanxi Province of China (grant number 2022JQ-552 and 2022JM-463), Undergraduate Training Program for Innovation and Entrepreneurship (121522015) and Scientific Research Fund for Health of Shaanxi Health Commission(2021D046).

Notes and references

1 E. M. Michalak, M. L. Burr, A. J. Bannister and M. A. Dawson, *Nat. Rev. Mol. Cell Biol.*, 2019, **20**, 573–589.
2 M. V. C. Greenberg and D. Bourc'his, *Nat. Rev. Mol. Cell Biol.*, 2019, **20**, 590–607.
3 Z. D. Smith and A. Meissner, *Nat. Rev. Genet.*, 2013, **14**, 204–220.
4 S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers and Y. Zhang, *Nature*, 2010, **466**, 1129–1133.
5 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.
6 M. Bachman, S. Uribe-Lewis, X. Yang, M. Williams, A. Murrell and S. Balasubramanian, *Nat. Chem.*, 2014, **6**, 1049–1055.
7 K. E. Szulwach, X. Li, Y. Li, C. X. Song, H. Wu, Q. Dai, H. Irier, A. K. Upadhyay, M. Gearing, A. I. Levey, A. Vasanthakumar, L. A. Godley, Q. Chang, X. Cheng, C. He and P. Jin, *Nat. Neurosci.*, 2011, **14**, 1607–1616.
8 M. Munzel, D. Globisch and T. Carell, *Angew. Chem., Int. Ed.*, 2011, **50**, 6460–6468.
9 P. Schuler and A. K. Miller, *Angew. Chem., Int. Ed.*, 2012, **51**, 10704–10707.
10 E. Kriukiene, Z. Liutkeviciute and S. Klimasauskas, *Chem. Soc. Rev.*, 2012, **41**, 6916–6930.
11 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.
12 H. Hu, M. Shu, L. He, X. Yu, X. Liu, Y. Lu, Y. Chen, X. Miao and X. Chen, *Br. J. Cancer*, 2017, **116**, 658–668.
13 Y. Tang, S. J. Zheng, C. B. Qi, Y. Q. Feng and B. F. Yuan, *Anal. Chem.*, 2015, **87**, 3445–3452.
14 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.
15 M. Mellen, P. Ayata, S. Dewell, S. Kriaucionis and N. Heintz, *Cell*, 2012, **151**, 1417–1430.
16 W. A. Pastor, U. J. Pape, Y. Huang, H. R. Henderson, R. Lister, M. Ko, E. M. McLoughlin, Y. Brudno, S. Mahapatra, P. Kapranov, M. Tahiliani, G. Q. Daley, X. S. Liu, J. R. Ecker, P. M. Milos, S. Agarwal and A. Rao, *Nature*, 2011, **473**, 394–397.
17 A. B. Robertson, J. A. Dahl, R. Ougland and A. Klungland, *Nat. Protoc.*, 2012, **7**, 340–350.
18 H. Stroud, S. Feng, S. Morey Kinney, S. Pradhan and S. E. Jacobsen, *Genome Biol.*, 2011, **12**, R54.
19 A. Szwagierczak, S. Bultmann, C. S. Schmidt, F. Spada and H. Leonhardt, *Nucleic Acids Res.*, 2010, **38**, e181.
20 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**, 1368–1380.
21 M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik and S. Balasubramanian, *Science*, 2012, **336**, 934–937.
22 Z. Zhang, X. Shan, P. Zhang, W. Liu, J. Yan and Z. Li, *Org. Biomol. Chem.*, 2019, **17**, 9849–9853.
23 Y. Du, Y. Lai, J. Y. Liu and J. Diao, *Small Methods*, 2021, **5**, e2100061.
24 Z. Zhang, J. Yan and Z. Li, *Chem. Commun.*, 2020, **56**, 3111–3114.
25 G. Hayashi, K. Koyama, H. Shiota, A. Kamio, T. Umeda, G. Nagae, H. Aburatani and A. Okamoto, *J. Am. Chem. Soc.*, 2016, **138**, 14178–14181.
26 K. Koyama, G. Hayashi, H. Ueda, S. Ota, G. Nagae, H. Aburatani and A. Okamoto, *Org. Biomol. Chem.*, 2021, **19**(29), 6478–6486.

27 A. Okamoto, K. Sugizaki, A. Nakamura, H. Yanagisawa and S. Ikeda, *Chem. Commun.*, 2011, **47**, 11231–11233.
28 N. Tomita, Y. Mori, H. Kanda and T. Notomi, *Nat. Protoc.*, 2008, **3**, 877–882.
29 T. Notomi, H. Okayama, H. Masubuchi, T. Yonekawa, K. Watanabe, N. Amino and T. Hase, *Nucleic Acids Res.*, 2000, **28**, E63.
30 Q. Xu, S. Q. Huang, F. Ma, B. Tang and C. Y. Zhang, *Anal. Chem.*, 2016, **88**, 2431–2439.
31 Y. Sun, B. Han and F. Sun, *RSC Adv.*, 2021, **11**, 17058–17063.
32 W. Du, M. Lv, J. Li, R. Yu and J. Jiang, *Chem. Commun.*, 2016, **52**, 12721–12724.
33 Z. Zhang, D. Yang, W. Tian, Y. Qi, W. Ren, Z. Li and C. Liu, *Anal. Chem.*, 2020, **92**, 3477–3482.
34 Y. Fu, X. Duan, J. Huang, L. Huang, L. Zhang, W. Cheng, S. Ding and X. Min, *Sci. Rep.*, 2019, **9**, 5955.