BIOLOGY Methods & Protocols

Biology Methods and Protocols, 2020, 1–9

doi: 10.1093/biomethods/bpaa019
Advance Access Publication Date: 28 September 2020
Methods Manuscript

METHODS MANUSCRIPT

Positive/negative ion-switching-based LC–MS/MS method for quantification of cytosine derivatives produced by the TET-family 5-methylcytosine dioxygenases

Aninda Sundar Dey ,‡ Navid J. Ayon,‡ Chayan Bhattacharya, William G. Gutheil, and Mridul Mukherji*

Division of Pharmacology & Pharmaceutical Sciences, University of Missouri-Kansas City, Kansas City, MO 64108, USA

*Correspondence address. Division of Pharmacology & Pharmaceutical Sciences, University of Missouri-Kansas City, MO 64108, USA. Tel: +1-816-235-6197; Fax: +1-816-235-5779; E-mail: mukherjim@umkc.edu

‡Equal contribution.

Abstract

Cytosine methylation at carbon-5 (5mC) in DNA plays crucial roles in epigenetic transcriptional regulation during metazoan development. The iron (II), 2-oxoglutarate-dependent Ten-Eleven Translocation (TET)-family dioxygenases initiate active demethylation of 5mC. TET2 oxidizes 5mC in nucleic acids into 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxylcytosine by iterative oxidation. Mutations in the TET2 gene are frequently detected in myeloid malignancies. Despite the established and emerging roles of TET oxygenases in health and diseases, *in vitro* characterization of these enzymes and their mutants is still in rudimentary stages. Here, we describe an improved positive/negative ion-switching-based liquid chromatography-tandem mass spectrometry (LC–MS/MS) method that can separate and quantify modified cytosine bases produced by TET-family 5-methylcytosine dioxygenases. This method will help in further elucidate the function of epigenetically important cytosine modifications. To the best of our knowledge, this is the first study reporting ion-switching-based LC–MS/MS method to analyse cytosine variants produced in TET catalysed reactions.

Keywords: epigenetics; DNA demethylation; TET dioxygenases; LC-MS/MS; positive/negative ion-switching

Introduction

The C5 position of cytosine within CpG dinucleotides is the predominant methylation site (5mCpG) in mammalian genomes [1]. In addition, a number of recent studies have uncovered extensive C5 cytosine methylation (5mC) in non-CpG sites (5mCpH, where H = A, T, or C) [2]. Methylation of cytosine at the C5 position is carried out by DNA methyl transferases, and mutations in these enzymes cause significant developmental defects [3]. Cytosine methylation upstream of a promoter region

Received: 26 August 2020; **Revised:** 15 September 2020. **Accepted:** 18 September 2020

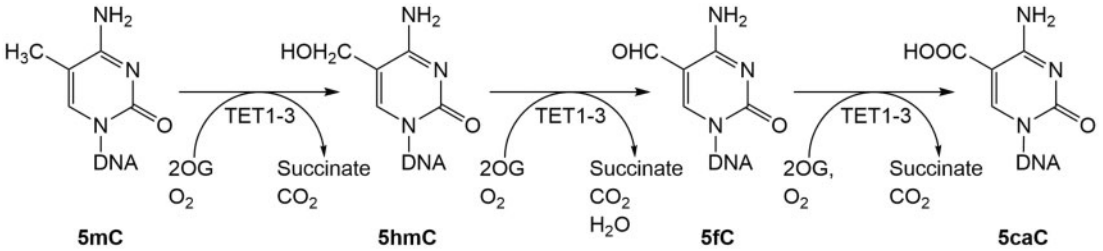

Figure 1: Schematic representation of sequential oxidation catalyzed by TET1-3 iron (II) and 2-oxoglutarate-dependent dioxygenases.

cause transcriptional repression of the downstream coding region and removal of the methylation marks restores transcription of the gene. The removal of 5mC marks are initiated by TET1-3 5mC dioxygenases [4]. These iron (II) and 2-oxoglutarate-dependent dioxygenases convert 5mC into 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) by sequential oxidation steps (Fig. 1). Finally, thymine-DNA glycosylase replaces 5fC and 5caC to unmodified cytosine residues using the base–excision repair pathway [5].

Due to its critical role in the epigenetic transcription regulation, 5mC and 5hmC are often termed as the fifth and sixth DNA bases (addition to the four normal bases). In the human genome, up to 4% of the cytosine residues are methylated at the C5 position. The 5mC modification generally serves as a transcriptional silencer at endogenous retrotransposons and gene promoters [6]. The 5mC modification also plays important roles in X chromosome inactivation, gene imprinting, nuclear reprogramming, and tissue-specific gene expression [7]. Unregulated 5mC modification is one of the fundamental features observed in many cancer types and several other diseases [8]. A number of methyl-CpG binding domain proteins have been identified, acting as readers of 5mC marks [9]. These readers further recruit chromatin remodelers, such as histone deacetylases resulting in transcription repression. In contrast to 5mC, genome-wide studies in embryonic stem (ES) cells and brain tissues demonstrated that 5hmC marks are enriched at or around transcriptional start sites, promoters with moderate to low CpG content and gene bodies, where its amount positively correlates to gene expression [10]. Recent studies suggest that several genes coding for proteins involved in the 5mC oxidation are often mutated in human tumors; and as a result, there is significant loss of 5hmC across many types of cancers [11]. Reduction in the 5hmC levels are also associated with the impaired self-renewal of ES cells [12]. Its stability and abundance would be compatible with the idea that 5hmC represents a true epigenetic mark which is recognized by specific readers such as Ubiquitin Like With PHD And Ring Finger Domains 2 which, in addition to 5mC, recognizes 5hmC marks [13].

For some time, 5fC and 5caC were considered as rare demethylation intermediates present in mammalian DNA. However, a recent study demonstrated 5fC as a stable DNA modification, suggesting that 5fC has functional roles in epigenetics [14]. This epigenetic mark is generally present at promoters of actively transcribed genes, transcriptional start sites, exons, and enhancers in ES cells [15, 16]. It appears that 5fC affects the rate of nucleotide incorporation and specificity of RNA polymerase II (RNAPII). The 5fC mark increases RNAPII backtracking and pausing, and reduces its fidelity of nucleotide incorporation [17]. The same studies by Kellinger *et al.* also suggest that 5caCs, like 5fC, have similar effects on RNAPII [17]. In addition, Yang *et al.* have demonstrated that the transcription factor TCF4 preferentially binds to DNA enriched in 5caC marks [18]. Taken together, these findings strongly suggest emerging important roles of cytosine derivatives/modifications (5mC, 5hmC, 5fC, and 5caC) in epigenetic transcription regulations.

As a result, substantial efforts have been made in developing advanced analytical methods to study nucleic acid modifications. These new methods have significantly contributed to our understanding of epigenetic modifications and their role on human genomics. Mass spectrometry has been widely used to study cytosine modifications due to its superior sensitivity, broad dynamic range, and ability to simultaneous quantification of these modifications. A number of groups, including ours, have reported liquid chromatography–tandem mass spectrometry (LC–MS/MS)-based analytical methods for quantification of cytosine derivatives [19–21]. However, it has been difficult to achieve optimum sensitivity for 5caC along with other three cytosine derivatives using a single LC–MS/MS method, possibly due to the difference in their polarity. In this study, we report a highly sensitive positive/negative ion-switching LC–MS/MS analytical method for separation and quantification of the cytosine derivatives produced by TET 5-methylcytosine dioxygenases. In order to achieve this, a number of solvent systems, pH conditions, and sample preparation procedures were evaluated to achieve the optimum analytical performance. The method described here will help further characterize the function of cytosine derivatives that play critical roles in health and diseases.

Materials and methods

All chemicals were from Thermo Fisher Scientific (Waltham, MA, USA) or Sigma-Aldrich (St Louis, MO, USA), unless otherwise stated, and were either analytical grade or higher. Nucleoside standards were purchased from Carbosynth (now Biosynth Carbosynth, Oakbrook Terrace, IL). *Escherichia coli* expression strain BL21 (DE3) was purchased from Novagen (now EMD Millipore, Billerica, MA, USA). All growth media were from Difco Laboratories (Detroit, MI, USA).

Optimization of MS/MS parameters for nucleosides

The LC–MS/MS analysis was performed on a Sciex 3200 QTrap mass spectrometer (Foster City, CA, USA) coupled to a Shimadzu UFLC LC-20 system (Columbia, MD, USA) using electrospray ionization (ESI) source and run with Analyst v 1.6.2 software. The C18 column (dimension 150 × 2 mm, particle size 5 μM, pore size 100 Å) for ion switching mode and C18 column (dimension 100 × 2 mm, particle size 5 μM, pore size 60 Å) for negative mode were obtained from keystone scientific (now Thermo Fisher Scientific). The UV detector on the HPLC was set at 280 nm, whereas the MS detections were performed in the

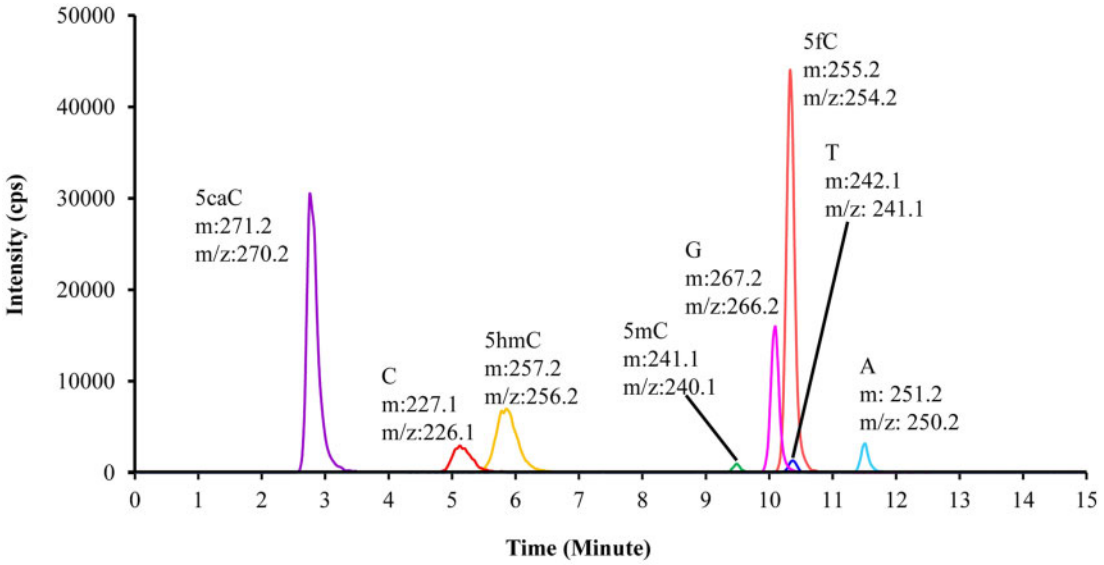

Figure 2: MRM chromatograms of eight different nucleosides in negative mode.

indicated ionization modes. Standard stock solutions of all eight nucleosides were prepared at a concentration of 100 μM in HPLC grade water. These stocks were used for optimization of various compound-dependent parameters on the mass spectrometer. Each nucleoside standard solution was infused into the mass spectrometer at a flow rate of 10 μl/min using a 4.6-mm Hamilton syringe (Franklin, MA, USA). Initial studies were performed in an enhanced mass-scan mode to detect the parent-ion peak under both positive and negative ionization mode, followed by generation of product-ion peaks in enhanced product-ion mode. Finally, the generated product-ions for each nucleoside were confirmed using enhanced precursor-ion mode. Various source-dependent parameters such as declustering potential, entrance potential, collision cell entrance potential, collision energy, and collision cell exit potential were optimized using the automated quantitative optimization feature of the Analyst software. Optimum parameters obtained from the infusion experiments were used to build a 2-min flow injection analysis method to optimize various source-dependent parameters. These parameters include curtain gas, temperature, gas flows, collision-associated-dissociation, and ion-spray voltage. These optimum parameters and the most intense fragment-ions per nucleoside in the positive, negative, and ion-switching modes in terms of sensitivity and resolution were used to build the final method.

Liquid chromatographic conditions for nucleosides under different MS/MS modes

A number of different solvent systems were evaluated for positive, negative, and ion-switching modes. These include water, ammonium acetate and ammonium formate as aqueous phases, and acetonitrile or methanol as organic phase. Effects of solvent composition and pH were also evaluated for optimum analytical performance. Specifically, for the positive ionization mode, our previously developed solvent system [where solvent A was 10-mM ammonium acetate pH 4.0 and solvent B was 20% acetonitrile/80% 10-mM ammonium acetate pH 4.0 (v/v)] was used [19]. For chromatographic separation of nucleosides in the negative ionization mode, ammonium acetate-based neutral solvent system was used where solvent A consists of 10-mM ammonium acetate pH 6.5 and solvent B consists of 80% acetonitrile/20% 10-mM ammonium acetate pH 6.5 (v/v). The gradient used was 0% solvent B (0-2 min), 0–20% solvent B (2–5 min), 20–60% solvent B (5–9 min), 60–0% solvent B (9–10 min), and with a 5-min post-equilibration with solvent A at a flow rate of 0.3 ml/min. For chromatographic separation of nucleosides in ion-switching mode, water/methanol-based solvent system was used where solvent A was water (adjusted to pH 3.5 using formic acid) and solvent B was methanol (adjusted to pH 3.0 using formic acid). The gradient used was 0% B (0–1 min), 0–2% B (1–12 min), 2–30% B (12–17 min), 30% B (17–18 min), 30–0% B (18–18.5 min), followed by a 4.5-min equilibration at 0% B at a flow rate of 0.3 ml/min.

LC–MS/MS analytical method validation

In order to construct the standard curves (and to determine the linearity) for all eight nucleosides in the positive, negative, and ion-switching mode, 100-μM standard stock solutions of each nucleoside were serially diluted in water. Diluted samples were analysed by the developed LC–MS/MS method (described above). Standard curves were drawn by plotting peak areas at different concentrations. For each nucleoside, the limit of detection (LOD) and lower limit of quantification (LLOQ) was calculated using (3.3× standard deviation of the lowest concentration)/slope of the standard curve and (10× standard deviation of the lowest concentration)/slope of the standard curve. To analyse the matrix effect (%ME), the peak area (cps) for all eight nucleosides were plotted at three different concentrations (0.2, 0.78, and 12.5 μM), and the slope of regression lines were calculated for both solvent and matrix. The matrix used for this study was comprised of all cofactors and TET2 enzyme in HEPES-NaOH buffer without the substrate DNA. The matrix effect was calculated according to the following formula, where 'S' is the slope of the linear regression curve:

$$\%ME = 100*(s_{matrix}/s_{water} - 1) \quad (1)$$

Purification of TET2 catalytic domain

Cloning and expression of the C-terminal TET2 dioxygenase domain (TET2 1129-1936, Δ1481-1843, the minimal catalytically active domain) were described in our earlier paper [19]. A bacterial pellet containing the recombinant TET2 protein was resuspended in 100-ml of 50 mM MOPS (3-(N-morpholino)propanesulfonic acid) buffer, pH 6.5 and sonicated for 5 × 30 s at power 20 with 60 s cooling intervals. The lysate was spun down at 4,000 r.p.m. for 45 min (Beckman Allegra X-15R centrifuge), and the soluble fraction was

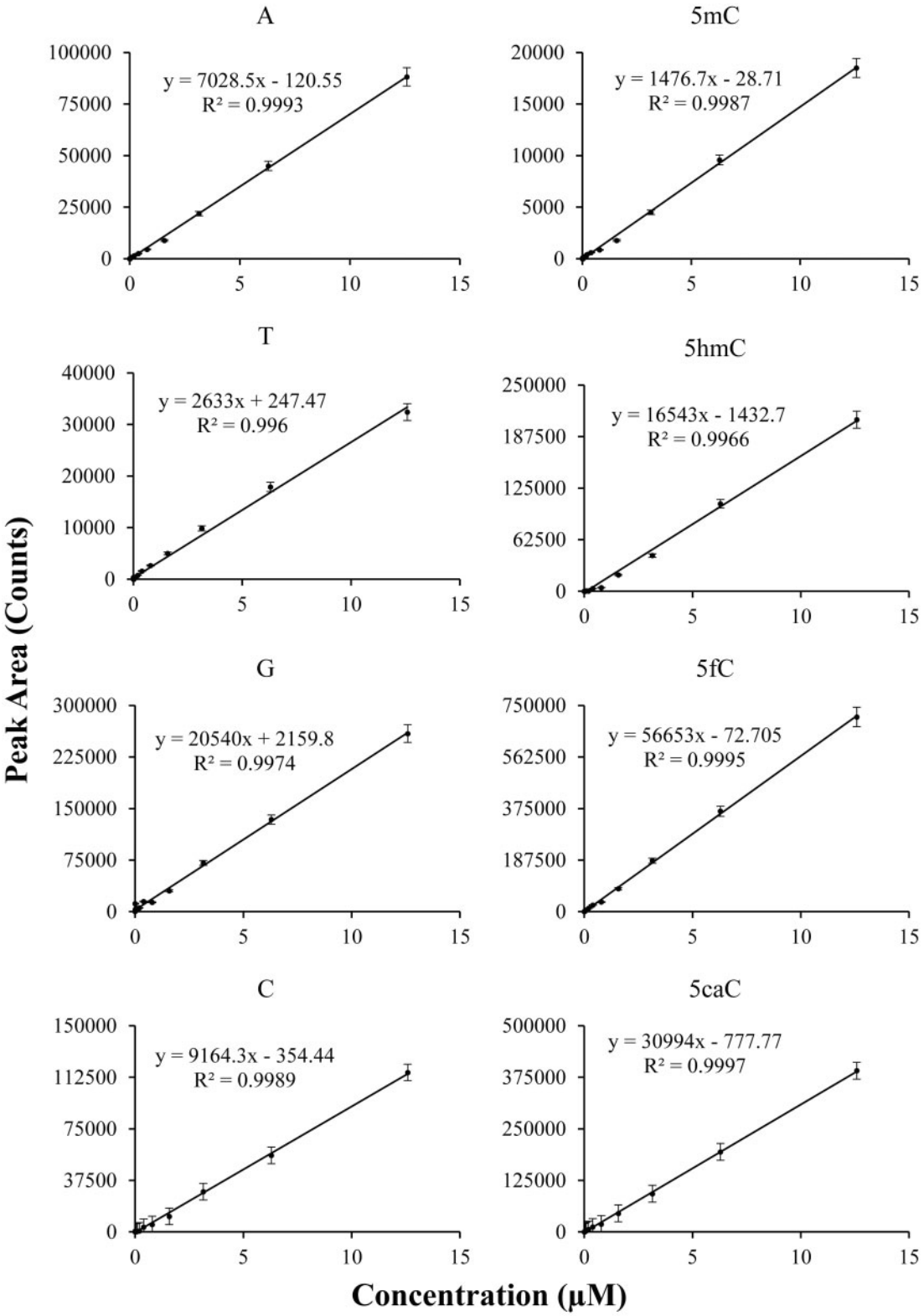

Figure 3: Standard curves of different nucleosides in negative mode.

filtered through a 0.45-μM filter before loading on an FPLC system. A strong cation exchange resin (10 ml of SP Sepharose) was packed into a XK26/20 FPLC column (Pharmacia now GE Healthcare, Piscataway, NJ, USA) and equilibrated in 10 bed volumes of wash buffer (50-mM MOPS buffer, pH 6.5) using an AKTA FPLC system (Pharmacia/GE Healthcare). Soluble fraction was loaded on the equilibrated column at 0.2 ml/min of flow rate. Following this, the column was washed with 5–10 bed volumes of wash buffer (flow rate 1 ml/min). Elution of the bound protein was carried out using a gradient (1 ml/min flow rate) from the wash buffer to the elution buffer (50-mM MOPS buffer, pH 6.5, 1 M NaCl). Zero to 100% of elution buffer was reached in 10 min, followed by holding at 100% elution buffer for 90 min, and then at 0% elution buffer for 30 min. Fractions containing purified TET2 were pooled and used for enzymatic analysis.

TET2 5-methylcytosine dioxygenase reaction

The TET2 enzymatic assay was performed using 25-mer double-stranded DNA (dsDNA) containing 5mC residues on both the strands as substrate (sense strand: 5′*AGCCCGCGCCG/iMe-dC*/5′.

Table 1: mass spectrometric parameters for the most intense MS/MS transitions of the eight nucleosides in negative mode[a]

Nucleosides	MRM Channels (Q1/Q3)	DP[b]	EP[c]	CEP[d]	CE[e]	CXP[f]	LOD (μM)	LLOQ (μM)
C	226.1/93.0	−35	−9	−10	−24	−3	0.10	0.32
5hmC	256.2/123.1	−50	−9	−14	−22	−3	0.20	0.65
5mC	240.1/124.1	−40	−10.5	−10	−16	−3	0.20	0.65
5caC	270.2/110.1	−50	−4	−12	−24	−3	0.05	0.16
T	241.2/125.1	−40	−8	−14	−12	−3	0.20	0.35
A	250.2/134.1	−35	−10	−10	−22	−3	0.10	0.32
5fC	254.2/121.1	−50	−2	−10	−24	−3	0.01	0.02
G	266.2/150.1	−60	−7.5	−14	−26	−3	0.003	0.01

[a]Global method parameters were: curtain gas 50, temperature 500°C, GS1 and GS2 (gas flows) 50 (arbitrary unit), collision-associated dissociation medium, ion spray voltage 4500 V for positive mode and −4500 V for negative mode.
[b]Declustering potential.
[c]Entrance potential.
[d]Collision cell entrance potential.
[e]Collision energy.
[f]Collision cell exit potential.

Results and discussion

LC–MS/MS-based analysis of nucleosides in negative mode

A number of methods have been described for the quantification of all four cytosine derivatives (5mC, 5hmC, 5fC, and 5caC) using LC–MS/MS with positive mode detection [19–21]. All these methods suffer from a common problem associated with low sensitivity for 5caC detection. We reasoned that this poor sensitivity for 5caC is due to difference in its polarity compared with other three cytosine derivatives under the LC–MS/MS ionization conditions. To address this issue, we tuned different mass spectrometry parameters for all eight nucleosides in the negative ionization mode. Using our previously developed solvent system (where solvent A was 10-mM ammonium acetate pH 4.0 and solvent B was 20% acetonitrile in 10-mM ammonium acetate pH 4.0) [19], 5caC was not detected in the negative mode with different gradient conditions attempted. However, at pH 6.5 using the slightly modified solvent system (solvent A: 10-mM ammonium acetate pH 6.5 and solvent B: 80% acetonitrile in 10-mM ammonium acetate pH 6.5), we were able to detect 5caC possibly because neutral to slightly basic pH is better for the detection of negatively charged species [22]. Thus, an LC–MS/MS method was developed for the separation and quantification for different cytosine nucleosides under negative ionization mode (see Materials and methods section). The gradient used for LC was able to resolve all eight nucleosides (Fig. 2). A standard curve for each nucleoside was drawn using serial dilutions of a 100-μM standard mixture containing all nucleosides (Fig. 3). For all analytes, standard curves were linear up to 1,000 pmol injection. Optimized MS parameters along with retention time, LOD, LLOQ, linearity of each analytes are listed in Table 1. Taken together, our new chromatographic condition in negative mode resulted in 32 times increased sensitivity for 5caC compared with a previously published paper [21]. Although the negative mode markedly improved the sensitivity for 5caC (and G), this approach resulted in lower sensitivities for six other nucleosides compared with the positive mode.

LC–MS/MS-based analysis of nucleosides in positive/negative ion-switching mode

Optimization process has demonstrated greater sensitivity for 5caC (and G to some extent) in the negative mode, whereas the

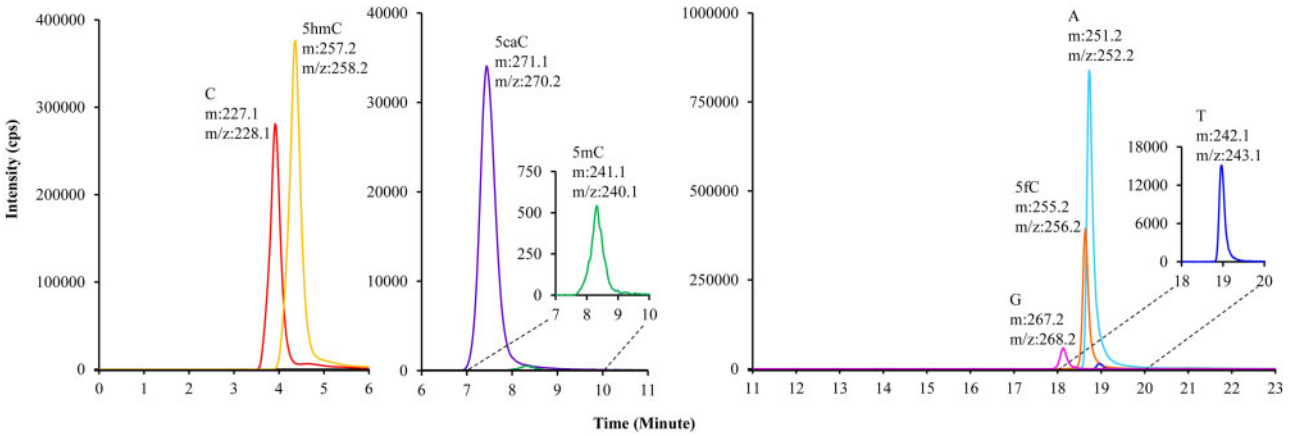

Figure 4: Ion-switching mode, MRM chromatograms of eight different nucleosides, period 1: positive mode (1–6 min); period 2: negative mode (6–11 min); period 3: positive mode (11–23 min).

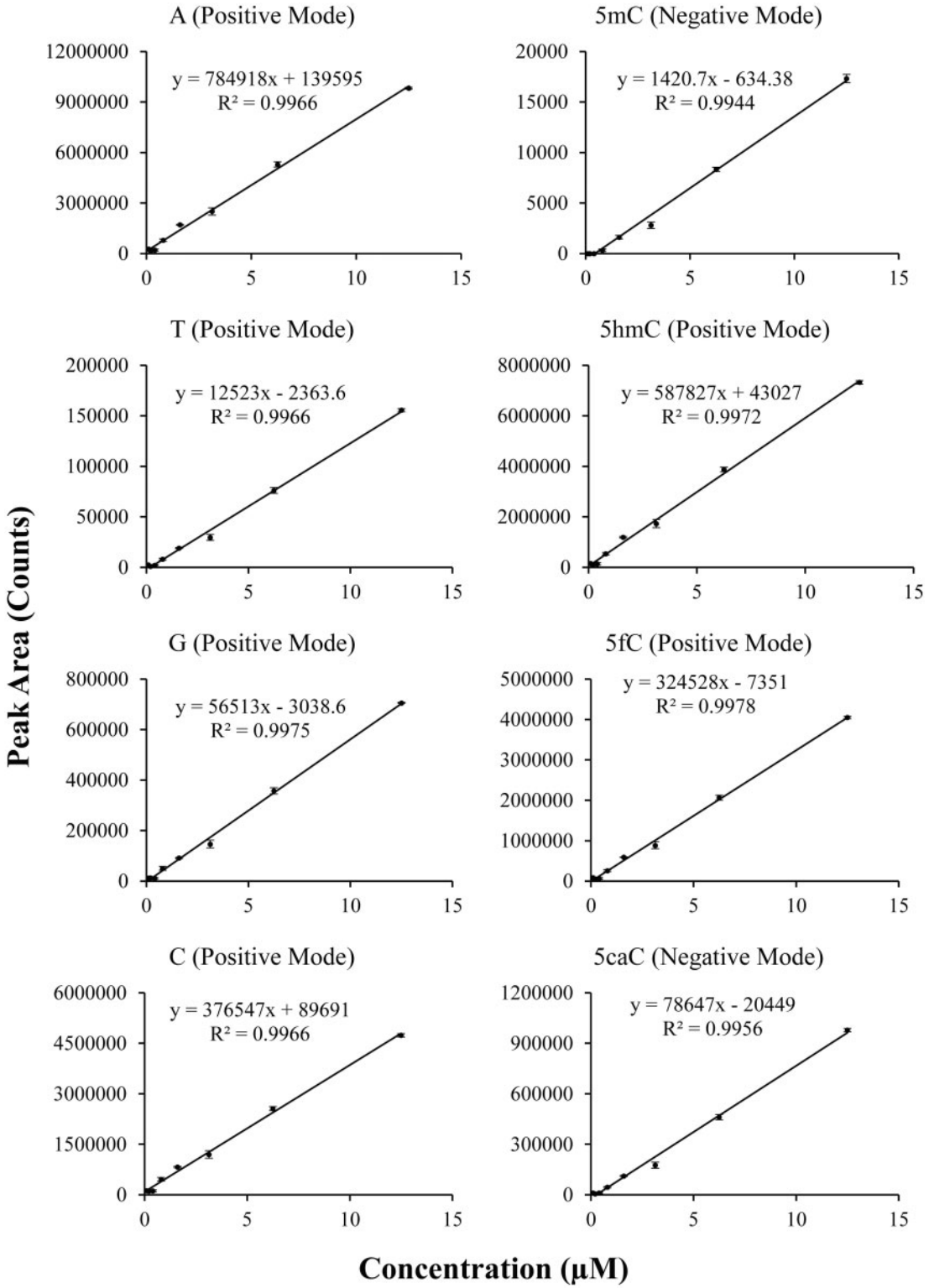

Figure 5: Standard curves of different nucleosides in ion-switching mode.

positive ion mode demonstrated higher sensitivities for six other nucleosides (A, T, C, 5mC, 5hmC, and 5fC). These results suggested that an ion-switching method incorporating periods of detection with superior sensitivity for specific nucleosides can lead to a highly sensitive LC–MS/MS analytical method. However, there were number of challenges that needed to be resolved in order to develop a mass spectrometry method with positive/negative ion switching. Since mass spectrometry in the positive and negative ionization mode, require different solvent systems, a common solvent system was required for both the modes under a single analytical run. Further, we observed that a minimum of 1–2 min is required for the mass spectrometer to

Table 2: Details of periods and mass spectrometric parameters for the most intense MS/MS transitions of the eight nucleosides in ion-switching mode[a]

Period	Mode	Duration (min)	Nucleosides	MRM channels (Q1/Q3)	DP[b]	EP[c]	CEP[d]	CE[e]	CXP[f]	LOD (μM)	LLOQ (μM)
1	Positive	0–6 (6)	C	228.1/112.1	21	7	14	15	4	0.36	1.09
			5hmC	258.2/142.1	36	3.5	14	17	4	0.25	0.75
2	Negative	6–11 (5)	5mC	240.1/124.1	−40	−10.5	−10	−16	−3	0.49	1.49
			5caC	270.2/110.1	−50	−4	−12	−24	−3	0.12	0.37
3	Positive	11–23 (12)	T	243.2/117.1	16	8	14	15	4	0.16	0.48
			A	252.2/136.1	41	9	14	17	4	0.33	1.00
			5fC	256.2/140.1	11	6	14	15	4	0.24	0.72
			G	268.2/152.1	21	7	14	37	4	0.25	0.76

[a]Global method parameters were: curtain gas 40, temperature 500°C, GS1 and GS2 (gas flows) 50 (arbitrary unit), collision-associated dissociation medium, ion spray voltage 4500 V for positive mode and −4500 V for negative mode.
[b]Declustering potential.
[c]Entrance potential.
[d]Collision cell entrance potential.
[e]Collision energy.
[f]Collision cell exit potential.

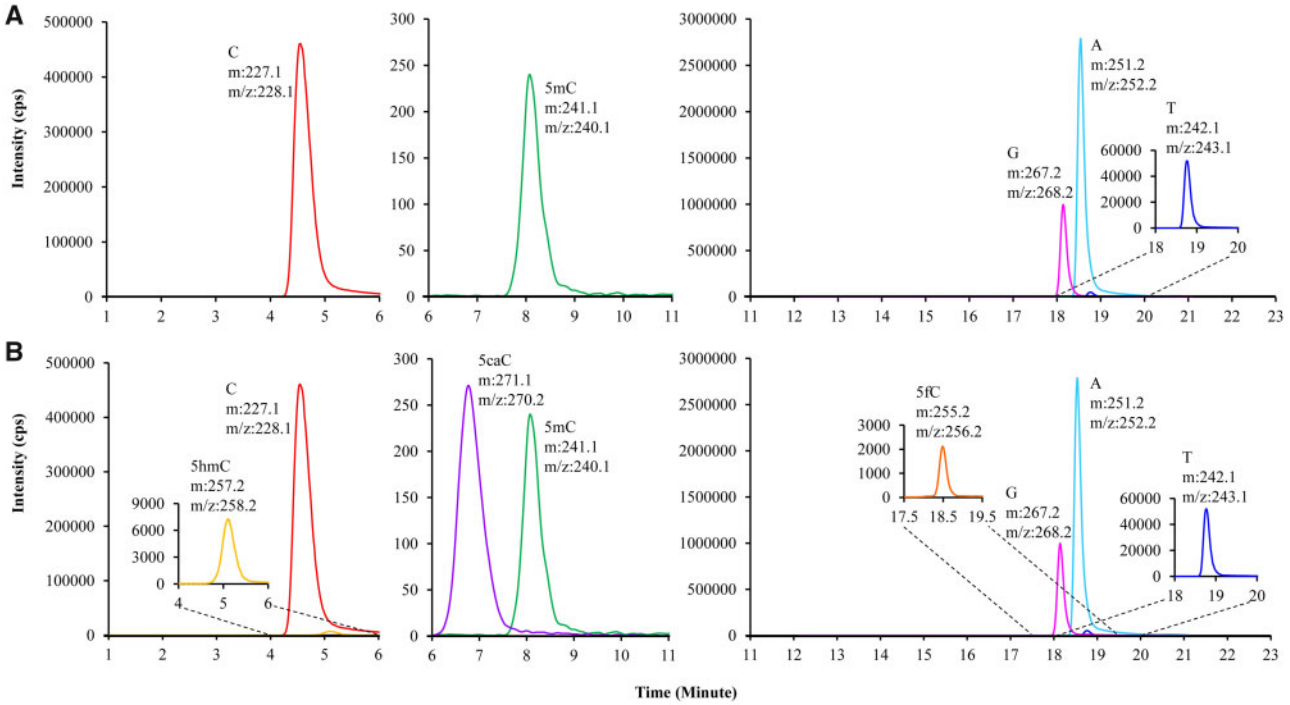

Figure 6: Ion-switching mode chromatograms of negative control assay, without the TET2 enzyme, indicating the presence of only A, T, G, C, and 5mC peaks (**A**). Ion-switching mode chromatograms of positive control assay, with TET2 enzyme, indicating the presence of three new peaks corresponding to 5hmC and 5fC in the positive mode and 5caC in the negative mode (**B**)

switch ionization modes to achieve the desired sensitivities and accommodate proper detection of the nucleosides within specific time window. Thus, the separation of all the nucleosides has to be done in such a way that the elution of 5caC (and its detection) can be performed in the negative mode period while for others (specifically 5hmC and 5fC) in the positive mode. This could be a challenge given similar properties of 5hmC and 5caC on a C18 column in reverse phase chromatography using most solvent systems [23].

In order to mitigate these issues, a number of solvent systems, pH, and chromatographic conditions were evaluated (data not shown). Finally, a water/methanol-based solvent system (described in Materials and methods section) at pH 3.5 separated 5caC from 5hmC and provided sufficient time for ion-switching and proper elution of the nucleosides in the determined time-window. Under these chromatographic conditions, C and 5hmC eluted first the under positive mode period followed by 5caC and 5mC, which eluted in the second negative mode period followed by the final period where A, T, G, and 5fC eluted under the positive mode (Fig. 4). Standard curves for each nucleoside were drawn using serial dilutions of a 100-μM standard mixture containing all nucleosides (Fig. 5). For all analytes, the standard curves were linear up to 1,000 pmol injection. Various optimized MS parameters along with retention time, LOD, LLOQ, linearity of each analytes are listed in Table 2. Although the sensitivity of 5caC at pH 3.5 (in water/methanol-

based solvent system) was little less compared with pH 6.5 (in ammonium acetate/acetonitrile-based solvent system), possibly due to the reason described earlier [22], sensitives for other nucleosides (particularly 5hmC and 5fC) in the positive mode increased several folds compared with ammonium acetate- or ammonium formate-based solvent systems. In comparison to our previously reported paper, sensitivity of 5hmC and 5fC improved by 6- to 8–folds, whereas 8- to 10-fold improvement achieved in case of 5caC. Thus, the developed ion-switching LC–MS/MS-based analytical method provides optimum sensitivity and resolution for all eight nucleosides.

Validation of positive/negative ion switching-based LC–MS/MS method using TET2 dioxygenase

The TET2 is one of the most frequently mutated genes in myelodysplastic syndromes (MDS) [24–26]. High degree of TET2 mutation is also observed in a number of myeloid malignancies such as MDS-myeloproliferative neoplasms (MDS-MPNs) and acute myeloid leukemia derived from MDS and MDS-MPN (sAML) [26]. So far, almost all studies evaluating TET2 enzymatic activity have relied on either (i) mass spectrometry-based assays or (ii) antibody-based quantification of cytosine derivatives produced by the TET2 enzyme. Limitations of the mass spectrometry-based assays are discussed earlier, whereas antibody-based assays (such as dot-blots [24] or ELISAs [27]) can quantify only one product (such as 5hmC or 5fC) at a time, thus not providing a full picture of the catalytic reaction carried by TET2.

Therefore, the activity of the TET2 catalytic domain was determined using a 25-mer dsDNA containing 5mC residue in a CpG island on both the strands as substrate. After TET2 enzymatic reactions, DNA oligonucleotides were purified and converted into nucleosides. The nucleoside mixtures were subjected to mass spectrometry analysis using the developed positive/negative ion-switching-based LC–MS/MS method. In the reactions without the TET2 enzyme (negative control), only A, T, G, C, and 5mC peaks were observed (Fig. 6A). However, in the positive control, which contained the TET2 dioxygenase, three new peaks corresponding to 5hmC and 5fC were observed in the positive mode and 5caC was observed in the negative mode (Fig. 6B).

Over the last decade, hundreds of frame-shift, nonsense, and missense mutations in the TET2 gene have been identified in patients [28]. However, lack in enzymatic characterizations have hampered our understanding of biochemical abnormalities associated with TET2 mutations. In this study, our results demonstrated that the developed positive/negative ion-switching-based LC–MS/MS method can be used to characterize the activity of the wild-type TET2 enzyme and its clinical mutants.

Funding

This research was funded by the US Department of Defense in the form of an Idea Award (W81XWH-13-1-0174), Aplastic Anemia & MDS Foundation Grant, and UMRB grant to M.M. W.G.G. was funded by NIH (R15-GM126502-01A1).

References

1. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* 2008;**9**:1–9.
2. Guo JU, Su Y, Shin JH *et al.* Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat Neurosci* 2014;**17**:215–22.
3. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 2007;**447**:425–32.
4. Tahiliani M, Koh KP, Shen Y *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 2009;**324**:930–5.
5. He Y-F, Li B-Z, Li Z *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 2011;**333**:1303–7.
6. Schultz MD, He Y, Whitaker JW *et al.* Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* 2015;**523**:212–6.
7. Feng S, Jacobsen SE, Reik W. Epigenetic reprogramming in plant and animal development. *Science* 2010;**330**:622–7.
8. Wajed SA, Laird PW, DeMeester TR. DNA methylation: an alternative pathway to cancer. *Ann Surg* 2001;**234**:10–20.
9. Du Q, Luu P-L, Stirzaker C *et al.* Methyl-CpG-binding domain proteins: readers of the epigenome. *Epigenomics* 2015;**7**:1051–73.
10. Koh KP, Rao A. DNA methylation and methylcytosine oxidation in cell fate decisions. *Curr Opin Cell Biol* 2013;**25**:152–61.
11. Pfeifer GP, Xiong W, Hahn MA *et al.* The role of 5-hydroxymethylcytosine in human cancer. *Cell Tissue Res* 2014;**356**:631–41.
12. Ye C, Li L. 5-Hydroxymethylcytosine: a new insight into epigenetics in cancer. *Cancer Biol Ther* 2014;**15**:10–5.
13. Chen R, Zhang Q, Duan X *et al.* The 5-hydroxymethylcytosine (5hmC) reader UHRF2 is required for normal levels of 5hmC in mouse adult brain and spatial learning and memory. *J Biol Chem* 2017;**292**:4533–43.
14. Bachman M, Uribe-Lewis S, Yang X *et al.* 5-Formylcytosine can be a stable DNA modification in mammals. *Nat Chem Biol* 2015;**11**:555–7.
15. Neri F, Incarnato D, Krepelova A *et al.* Single-base resolution analysis of 5-formyl and 5-carboxyl cytosine reveals promoter DNA methylation dynamics. *Cell Rep* 2015;**10**:674–83.
16. Song C-X, Szulwach KE, Dai Q *et al.* Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell* 2013;**153**:678–91.
17. Kellinger MW, Song C-X, Chong J *et al.* 5-Formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat Struct Mol Biol* 2012;**19**:831–3.
18. Yang J, Horton JR, Li J *et al.* Structural basis for preferential binding of human TCF4 to DNA containing 5-carboxylcytosine. *Nucleic Acids Res* 2019;**47**:8375–87.
19. Bhattacharya C *et al.* Efficient purification and LC-MS/MS-based assay development for ten-eleven translocation-2 5-methylcytosine dioxygenase. *J Vis Exp* 2018;**140**:57798.
20. Jaiswal M, Bhar S, Vemula H *et al.* Convenient expression, purification and quantitative liquid chromatography-tandem mass spectrometry-based analysis of TET2 5-methylcytosine demethylase. *Protein Expr Purif* 2017;**132**:143–51.
21. Hu L, Li Z, Cheng J *et al.* Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell* 2013;**155**:1545–55.
22. Ayon NJ, Sharma AD, Gutheil WG. LC-MS/MS-based separation and quantification of Marfey's reagent derivatized

proteinogenic amino acid DL-stereoisomers. *J Am Soc Mass Spectrom* 2019;**30**:448–58.
23. Roberts CE, Raner GM, Isaacs GD. High performance liquid chromatography separation of epigenetic cytosine variants. *Methods Protoc* 2018;**1**:10.
24. Ko M, Huang Y, Jankowska AM *et al*. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* 2010;**468**:839–43.
25. Langemeijer SMC, Kuiper RP, Berends M *et al*. Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat Genet* 2009;**41**:838–42.
26. Smith AE, Mohamedali AM, Kulasekararaj A *et al*. Next-generation sequencing of the TET2 gene in 355 MDS and CMML patients reveals low-abundance mutant clones with early origins, but indicates no definite prognostic value. *Blood* 2010; **116**:3923–32.
27. Blaschke K, Ebata KT, Karimi MM *et al*. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* 2013;**500**:222–6.
28. Ponnaluri VK, Maciejewski JP, Mukherji M. A mechanistic overview of TET-mediated 5-methylcytosine oxidation. *Biochem Biophys Res Commun* 2013;**436**:115–20.