Title

Normalized Retention Time for Targeted Analysis of the DNA Adductome.

Journal

Analytical Chemistry, 90(24)

Authors

Cui, Yuxiang
Wang, Pengcheng
Yu, Yang
et al.

Publication Date

2018-12-18

DOI

10.1021/acs.analchem.8b04660

Peer reviewed

Published in final edited form as:

Anal Chem. 2018 December 18; 90(24): 14111–14115. doi:10.1021/acs.analchem.8b04660.

Normalized Retention Time for Targeted Analysis of the DNA Adductome

Yuxiang Cui[†], **Pengcheng Wang**[‡], **Yang Yu**[†], **Jun Yuan**[†], and **Yinsheng Wang**[*,†,‡]

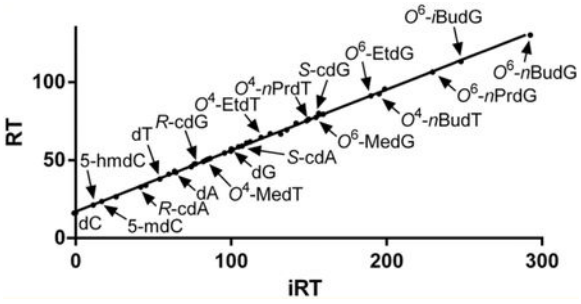
[†]Environmental Toxicology Graduate Program, University of California, Riverside, California 92521, United States

[‡]Department of Chemistry, University of California, Riverside, California 92521, United States

Abstract

A wide spectrum of DNA lesions can be generated from byproducts of endogenous metabolism and/or from environmental exposure. A DNA adductomic approach for the robust quantification of DNA adducts in cellular and tissue DNA may facilitate the use of DNA adducts for biomonitoring studies and enable comprehensive assessment about DNA repair. Normalized retention time (iRT) has been widely used in scheduled selected-reaction monitoring (SRM) methods for highly sensitive and high-throughput analyses of protein samples in complicated matrices. By using a similar method, we established the iRT scores for 36 modified nucleosides from the retention times of the four canonical 2′-deoxynucleosides on a nanoflow liquid chromatography–nanospray ionization–tandem mass spectrometry (nLC–NSI–MS/MS) system. The iRT scores facilitated reliable prediction of retention time and were employed for establishing a scheduled SRM method for quantitative assessment of a subset of the DNA adductome. The quantification results of the scheduled SRM method were more accurate and precise than those from an unscheduled method.

Graphical Abstract:



DNA damage constantly arises from environmental exposure and endogenous metabolism. Humans are frequently exposed to many DNA-damaging agents simultaneously, which give rise to a complicated array of DNA lesions. For instance, tobacco smoking and alcohol consumption can both lead to the formation of multiple DNA lesions.[1,2] If not properly

[*]**Corresponding Author**, Phone: (951)827-2700. yinsheng.wang@ucr.edu.

Notes

The authors declare no competing financial interest.

repaired, DNA lesions may confer cytotoxic or mutagenic consequences. Hence, the assessment of the occurrence and repair of DNA lesions is important for understanding the human health consequences of DNA damage. Such assessment can be facilitated by comprehensive measurement of the DNA adductome.

Liquid chromatography–tandem mass spectrometry (LC–MS/MS) has been routinely employed in proteomics and metabolomics, and it has also been recently used in DNA adductomics.[3] Previously published LC–MS/MS approaches for DNA adductomic analysis include selected-reaction monitoring (SRM),[4,5] high-resolution full-scan MS,[6] high-resolution data-dependent MS^2 acquisition and MS^3 fragmentation triggered by constant neutral-loss scanning,[7,8] and data-independent acquisition using high-resolution MS/MS in selected-ion monitoring mode with a wide *m/z* window for precursor ion isolation.[9] Data-dependent approaches usually harness the characteristic loss of a deoxyribose or nucleobase in MS^2. Such approaches are capable of monitoring known and unknown adducts, though adducts of low abundance may escape detection. In addition, some lesions do not undergo fragmentation via neutral loss of a deoxyribose, e.g., 8,5′-cyclopurine-2′-deoxynucleoside lesions.[10] In contrast, targeted approaches such as SRM can offer better specificity and sensitivity. A high-throughput and comprehensive adductomic approach requires the inclusion of as many lesions into the analytical workflow as possible. With SRM, a longer transition list requires longer cycle time, which leads to smaller number of data points for defining a chromatographic peak that may diminish the measurement precision and accuracy. To increase the number of analytes that can be monitored in a single LC–MS/MS experiment, acquisition can be limited to a few minutes around the retention time (RT) by scheduled SRM.[11] Recent developments in scheduled SRM analysis of peptides and proteins have led to targeted and quantitative analytical methods for samples in complex matrices. In this regard, normalized retention time (iRT) has been widely used for scheduled SRM analyses of peptides, where iRT is empirically derived from the RT of analyte of interest relative to those of standard peptides.[12] The iRT scores are specific and stable across different chromatographic configurations and conditions. Inspired by these approaches employed in proteomic studies, herein we incorporated iRT and developed a scheduled SRM method for DNA adductome analysis.

We measured empirical RT with a Dionex UltiMate 3000 RSLCnano system coupled with a Q Exactive Plus quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific), using an in-house packed trapping column (porous graphitic carbon (PGC), 5 μm in particle size, Thermo Fisher Scientific) and an in-house packed analytical column (Zorbax SB-C18, 5 μm in particle size, 100 Å in pore size, Agilent), where mobile phases A and B were 0.1% (v/v) formic acid in water and 0.1% (v/v) formic acid in acetonitrile/water (4:1, v/v), respectively. In proteomic applications, a group of standard peptides are used as references for retention time calibration. To employ iRT in DNA adductomic analysis, we chose the 4 canonical 2′-deoxynucleosides, i.e., 2′-deoxycytidine (dC), thymidine (dT), 2′-deoxyadenosine (dA), and 2′-deoxyguanosine (dG), as references. We arbitrarily assign the iRT scores of 0 and 100 to the earliest (dC) and latest (dG) eluting nucleosides, respectively. iRT scores for dT and dA were calculated from their RT values by linear regression to dC and dG using the previously described method.[12]

The iRT scores of the 4 canonical nucleosides were then applied to a linear gradient of 10–75% B in 5–150 min (hereafter referred to as “long and fast gradient”) to establish iRT scores for modified nucleosides. We subsequently measured the RT and established the iRT scores for 36 modified nucleosides (SI Materials), including alkylation adducts, oxidatively induced DNA adducts, and several epigenetic modifications (Figure 1, and Figure S1 and Table S1). The iRT scores for 4 canonical RNA nucleosides, i.e., cytidine (rC), uridine (U), adenosine (rA), guanosine (rG), were also obtained considering that residual amount of RNA may be present in DNA samples. These iRT scores can be conveniently shared among different laboratories since the reference canonical nucleosides can be readily obtained and are usually ubiquitous if nucleoside mixtures of cellular or tissue DNA are used for LC–MS and MS/MS analyses.

To determine if the RT–iRT correlation is tolerant of changes in chromatographic conditions, we tested four different gradients: a short gradient (linear at 5–35 min 10–50% B), a short and fast gradient (linear at 5–35 min 10–75% B), a long gradient (linear at 5–150 min 10–50% B), and a long and fast gradient (linear at 5–150 min 10–75% B). The short gradient led to the elution of fewest targeted analytes in the linear portion of the gradient, whereas the long and fast gradient resulted in the elution of most targeted analytes. Although fewer targeted nucleosides were eluted with the short gradient, the RT–iRT correlation for those analytes that elute during the linear gradient exhibited good linearity under different gradients (short gradient, $R^2 = 0.990$; short and fast gradient, $R^2 = 0.975$; long gradient, $R^2 = 0.997$; long and fast gradient, $R^2 = 0.998$) (Figure 2). With the same stationary phase material and mobile phase compositions, iRT also exhibited an excellent linear correlation with RT on another LC platform (i.e., an EASY-nLC1200 system, Figure S2).

With short gradient, dG (iRT = 100) elutes near the end of linear gradient (~35 min). As a result, many targeted species with predicted RT being later than 35 min do not follow the same linear correlation with the canonical nucleosides. With a longer gradient, especially the long and fast gradient, dG elutes in the midst of linear gradient (~57 min), yet the RT–iRT correlation is still linear in the retention time range after the elution of dG. When iRT values are used to predict RT, the analytes of interest should fall on the same linear gradient as the reference nucleosides. If target nucleosides do not elute during the linear gradient, employing other nucleosides as references and linear fit between neighboring reference nucleosides should allow for robust retention time prediction.

The elution of analytes depends on both the length of the gradient and the highest percentage of mobile phase B at the end of the linear gradient. For gradients of the same length, a higher %B allows for the elution of more targeted analytes. Additionally, when the gradients start and end at the same %B, a longer gradient facilitates the elution of more targeted analytes. The %B where an analyte elutes does not remain constant when the gradient changes; nevertheless, the order of elution does not alter on the same type of columns. On the grounds that iRT scores are normalized by reference analytes, iRT-based RT prediction can overcome better RT drift induced by changes in chromatographic conditions.

The RT and iRT for these nucleosides also displayed an excellent linear correlation when a different type of stationary phase material was used for the analytical column (Magic C18

AQ, Michrom BioResources) (Figure 3, Figure S3). In addition, the iRTs obtained on the Magic C18 AQ columns are very similar to those obtained on the aforementioned Zorbax SB-C18 columns (Figure 3, Figure S4, Table S1), suggesting that iRT values do not vary substantially with different types of reversed-phase C18 materials.

To assess whether iRT can help improve quantification, we compared the empirical RT with RT predicted from iRT with the combination of a PGC trapping column and a Zorbax SB-C18 analytical column on the EASY-nLC 1200 system. Our results showed that the absolute differences between the predicted RT and empirical RT are within 3 min for most nucleosides (Figure 4a). Therefore, we implemented an acquisition window of 6 min for scheduled SRM experiment. For a few nucleosides with iRT scores being much larger than 100, e.g., O^6-*n*-propyl-2′-deoxyguanosine (O^6-*n*PrdG) and O^6-*n*-butyl-2′-deoxyguanosine (O^6-*n*BudG), when their RT was projected on a longer gradient (i.e., long gradient, and long and fast gradient), the difference between predicted and observed RT is larger (Figure 4a), which may require a wider acquisition window (e.g., 12 min) for these nucleosides. Considering that few targeted species elute during the latter part of the gradient, such a wide acquisition window would not appreciably compromise quantification results.

Next, we compared the quantification results acquired from the scheduled SRM method with those obtained via traditional quantification method and unscheduled adductomic method. In this regard, stable-isotope dilution has been widely used for the quantification of DNA adducts.[13] Isotopically labeled standards exhibit the same chromatographic behavior as their corresponding unlabeled analytes, though they differ in mass. Conventionally, stable-isotope dilution coupled with LC–MS/MS methods only involve monitoring a few scan events for one or two analytes of interest; thus, the cycle time is short enough to enable robust quantification. An adductomic approach-based method requires monitoring a large number of lesions. Quantitation accuracy and precision may decrease as cycle time increases with the increase in the number of scan events.

We chose two oxidatively induced DNA lesions for which we have established robust quantification methods,[10] i.e., (5′*R*)-8,5′-cyclo-2′-deoxyguanosine (*R*-cdG) and (5′*R*)-8,5′-cyclo-2′-deoxyadenosine (*R*-cdA), to evaluate the quantification performance of the scheduled SRM-based LC–MS/MS method. The expected peak area ratios (i.e., 100% recovery) of unlabeled and isotopically labeled nucleosides were defined as what would be obtained with a traditional targeted quantification method (SI Materials) (hereafter referred to as “cPu method”) with the inclusion of only four precursor ions. In the unscheduled adductomic method (SI Materials), 44 precursors were monitored throughout the gradient. In the scheduled adductomic method (SI Materials), an acquisition window of 6 min was imposed for each nucleoside, where 4 and 8 precursor ions were simultaneously monitored in the acquisition windows of *R*-cdA and *R*-cdG, respectively. The reduction in number of scan events resulted in a more accurate and precise quantification relative to the unscheduled adductomics method (Figure 4b). We normalized the peak area ratios to the mean of peak area ratios obtained with the cPu method. In this way, the mean value of each method represents the quantification accuracy (i.e., percent recovery). Compared with the cPu method, which yields a relative standard deviation (RSD) of 1.9% (*R*-cdA) and 6.0% (*R*-cdG), the unscheduled adductomics method gives RSD of 5.0% (*R*-cdA) and 13.6% (*R*-

cdG), recovery of 98.3% (*R*-cdA) and 97.8% (*R*-cdG), while the scheduled adductomics method provides a RSD of 0.8% (*R*-cdA) and 5.3% (*R*-cdG), recovery of 99.1% (*R*-cdA) and 100.6% (*R*-cdG). Quantification results provided by the scheduled adductomics method are comparable to the cPu method. Hence, scheduled SRM can significantly increase the number of DNA lesions that could be quantitatively measured in a single LC–MS/MS experiment.

DNA adductome analysis can be challenging in that (1) many regioisomers may exist; (2) the hydrophobicity differs greatly; and (3) the adduct levels are very low. Although acid and thermal hydrolysis are sometimes used to release adducts from DNA,[3,13] DNA is often first enzymatically digested into nucleosides for LC–MS/MS analysis of DNA adducts. In the positive-ion mode, collision-induced dissociation of these nucleosides led to the neutral loss of a 2-deoxyribose, which does not provide diagnostic MS^2 for regioisomers. Employing MS^3 on a linear ion-trap mass spectrometer gives more structural information for adduct identification, yet some regioisomers still have identical MS^3 spectra. Meanwhile, with LC separation, the RT of an analyte can provide another dimension of specificity, thereby allowing unambiguous identification when in combination with high-resolution MS^2 spectrum. Modified nucleosides have very different hydrophobicity; for example, 5-hydroxymethyl-2′-deoxycytidine (5-hmdC) is much more hydrophilic than those DNA adducts induced by metabolites of polycyclic aromatic hydrocarbons. With the gradients employed in this work, we were able to cover many relatively hydrophilic modifications including 5-methyl-2′-deoxycytidine (5-mdC) and 5-hmdC, and several hydrophobic adducts, e.g., the O^6-alkyl-2′-deoxyguanosine (O^6-alkyldG) lesions. Incorporation of more hydrophobic adducts, such as those induced by polycyclic aromatic hydrocarbons and heterocyclic aromatic amines, may eventually be achieved by using different stationary phase materials and different reference nucleosides for establishing the RT–iRT relationship. Considering the low abundance of DNA lesions (usually a few per 10^{10} nucleosides to a few per 10^6 nucleosides),[13] removal of canonical nucleosides with off-line HPLC or other approaches can prevent the overloading of HPLC columns.[14] Regardless of the possibility of missing unknown adducts, a targeted adductomics approach that takes both RT and MS information into consideration is more accurate in identifying modified nucleosides. This would be especially helpful when high-resolution mass spectrometers are not available. For instance, 3,N^4-etheno-2′-deoxycytidine (edC) and dA would have the same MS/MS transition in the positive-ion mode (m/z 252 → 136). Retention time is unarguably an important factor for determining the identity of an adduct. Future work should emphasize on expanding the method to include as many adducts as possible.

To conclude, we applied, for the first time, the concept of iRT into the targeted analysis of the DNA lesions in an adductomic approach. We established the iRT scores for 36 modified 2′-deoxynucleosides and 4 canonical RNA nucleosides based on linear regression analysis of experimentally determined RT values for these lesions and the canonical unmodified nucleosides. These iRT scores can be used for predicting RT in new chromatographic settings, thereby allowing for targeted quantification of DNA adducts. Although standards are required for establishing iRT scores, these iRT scores can be easily transferred to new chromatographic setups when using columns packed with similar stationary phase materials.

It can be envisaged that the method can be generally applicable for the comprehensive investigation of the occurrence and repair of DNA lesions in the future.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (R01 ES029749).

REFERENCES

(1). Brooks PJ; Theruvathu JA Alcohol 2005, 35, 187–193. [PubMed: 16054980]
(2). Phillips DH Carcinogenesis 2002, 23, 1979–2004. [PubMed: 12507921]
(3). Villalta PW; Balbo S Int. J. Mol. Sci 2017, 18, 1870.
(4). Kanaly RA; Hanaoka T; Sugimura H; Toda H; Matsui S; Matsuda T Antioxid. Redox Signaling 2006, 8, 993–1001.
(5). Matsuda T; Tao H; Goto M; Yamada H; Suzuki M; Wu YJ; Xiao N; He Q; Guo WW; Cai ZM; Kurabe N; Ishino K; Matsushima Y; Shinmura K; Konno H; Maekawa M; Wang YP; Sugimura H Carcinogenesis 2013, 34, 121–127. [PubMed: 23066087]
(6). Hemeryck LY; Decloedt AI; Vanden Bussche J; Geboes KP; Vanhaecke L Anal. Chim. Acta 2015, 892, 123–131. [PubMed: 26388482]
(7). Balbo S; Hecht SS; Upadhyaya P; Villalta PW Anal. Chem 2014, 86, 1744–1752. [PubMed: 24410521]
(8). Stornetta A; Villalta PW; Hecht SS; Sturla SJ; Balbo S Anal. Chem 2015, 87, 11706–11713. [PubMed: 26509677]
(9). Guo J; Villalta PW; Turesky RJ Anal. Chem 2017, 89, 11728–11736. [PubMed: 28977750]
(10). Wang J; Yuan BF; Guerrero C; Bahde R; Gupta S; Wang YS Anal. Chem 2011, 83, 2201–2209. [PubMed: 21323344]
(11). Lange V; Picotti P; Domon B; Aebersold R Mol. Syst. Biol 2008, 4, 222. [PubMed: 18854821]
(12). Escher C; Reiter L; MacLean B; Ossola R; Herzog F; Chilton J; MacCoss MJ; Rinner O Proteomics 2012, 12, 1111–1121. [PubMed: 22577012]
(13). Tretyakova N; Goggin M; Sangaraju D; Janis G Chem. Res. Toxicol 2012, 25, 2007–2035. [PubMed: 22827593]
(14). Hong HZ; Cao HC; Wang YS; Wang YS Chem. Res. Toxicol 2006, 19, 614–621. [PubMed: 16696563]

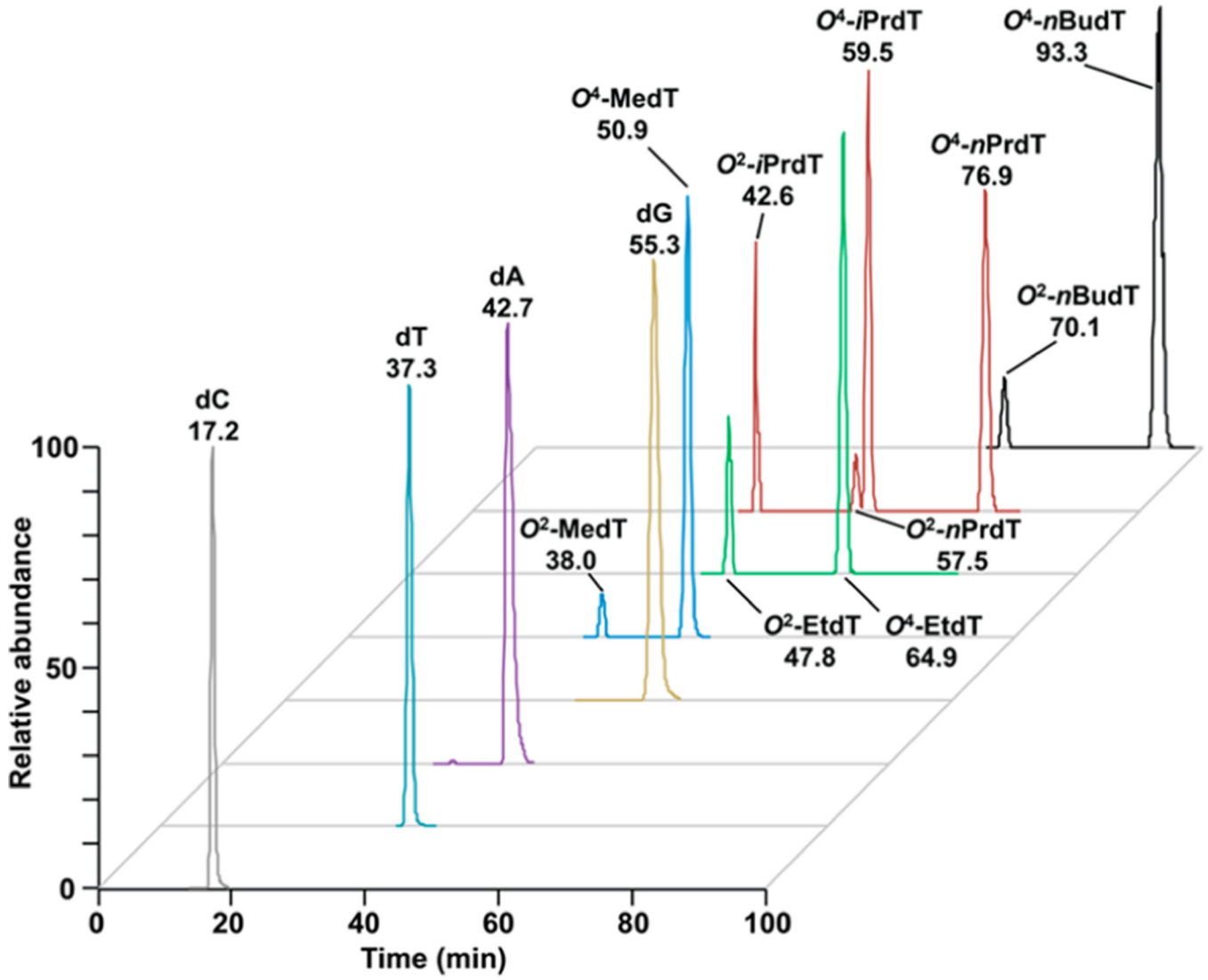


Figure 1.

Representative SICs from a scheduled SRM-based LC–MS/MS experiment. Chromatographic peaks for canonical 2′-deoxynucleosides, O^2-alkylthymidine (O^2-alkyldT), O^4-alkylthymidine (O^4-alkyldT), and their retention times are shown, where the alkyl groups are —CH_3 (—Me), —CH_2CH_3 (—Et), —$(CH_2)_2CH_3$ (—*n*Pr), —$CH(CH_3)_2$ (—*iPr*), and —$(CH_2)_3CH_3$ (—*n*Bu).

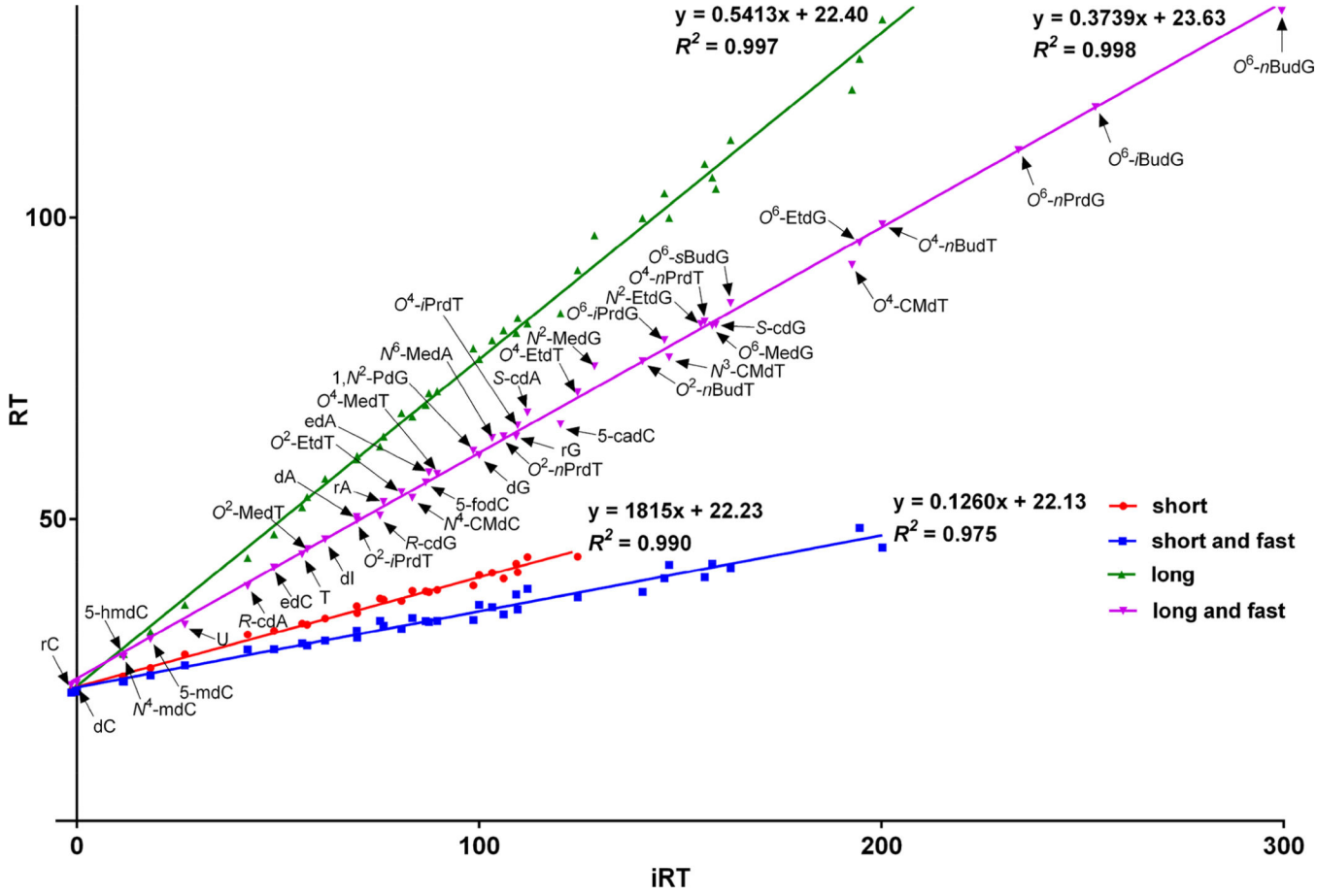


Figure 2.

Correlation between RT and iRT obtained on a porous graphitic carbon (PGC) trapping column and a Zorbax SB-C18 analytical column, with different gradients. Nucleosides are labeled on the line from results obtained from the use of a long and fast linear gradient (0–75% B in 5–150 min), and the complete names of DNA lesions are listed in the Supporting Information.

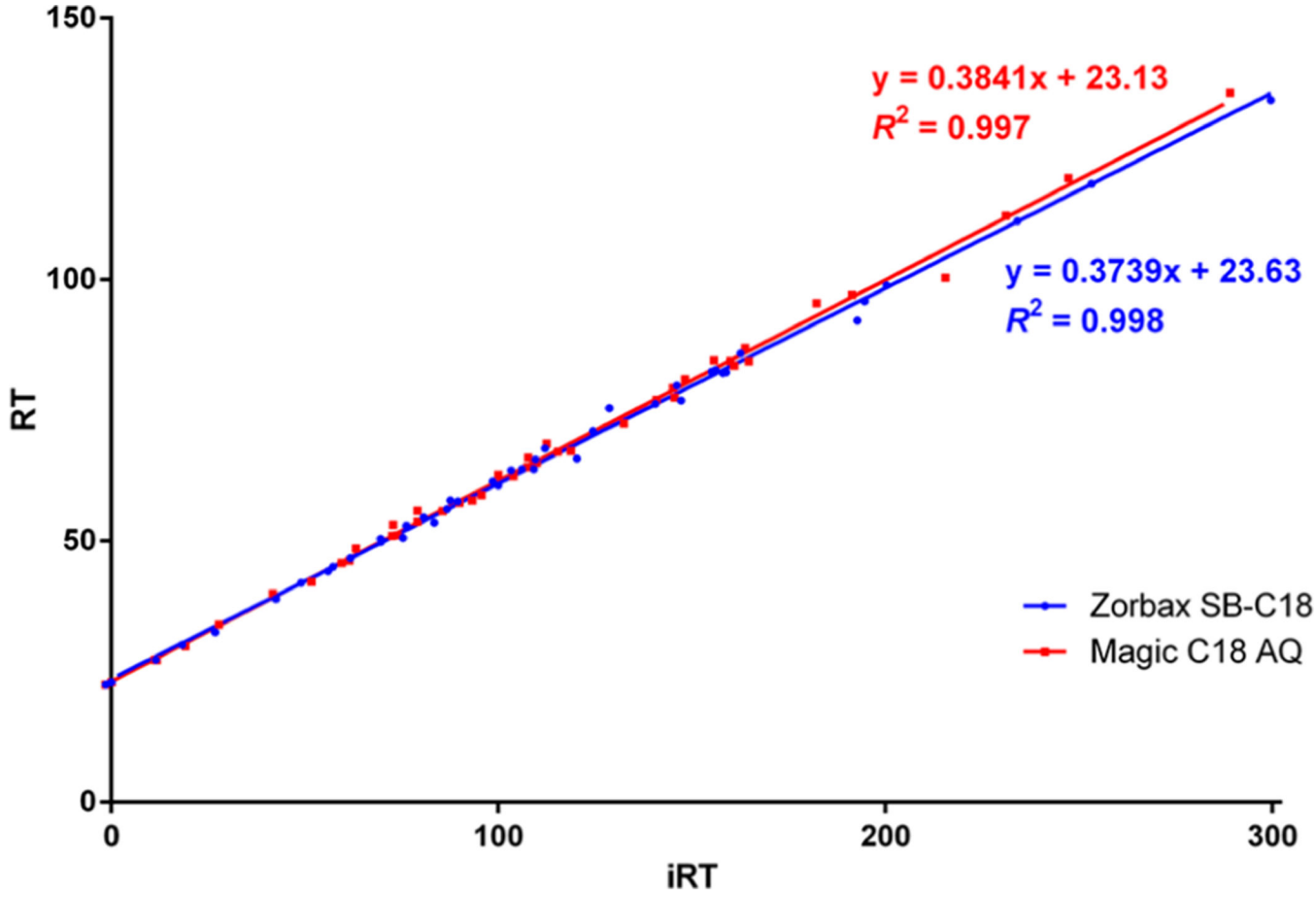


Figure 3.
RT–iRT correlation on Magic C18 AQ columns and that on Zorbax SB-C18 columns.

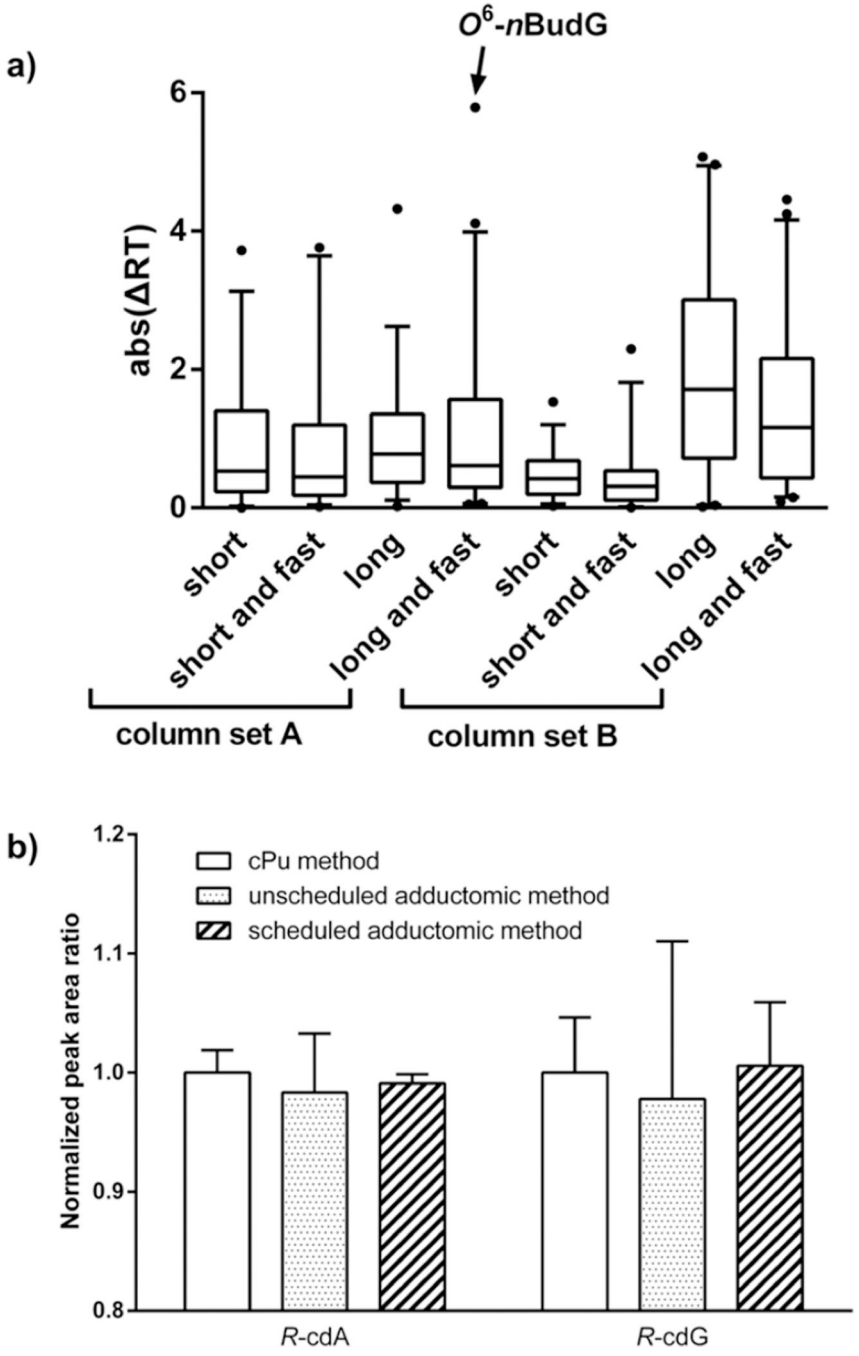


Figure 4.

(a) Absolute difference (in min) between RT predicted from iRT and empirical RT with different HPLC gradients. Two sets of columns with the same stationary phase materials (PGC trapping columns and Zorbax SB-C18 analytical columns), but with different lengths, were compared (column set A, trapping column 3.1 cm, analytical column 25.1 cm; column set B, trapping column 2.8 cm, analytical column 25.8 cm). The differences between observed/empirical RT and iRT-predicted RT for most nucleosides were within 3 min. A typical outlier, O^6-nBudG, was marked on the graph. Whiskers were plotted from the 5th to

95th percentile. (b) Accuracy and precision of the unscheduled and scheduled adductomic method compared with conventional quantification method. cPu method: conventional quantification method monitoring 4 precursor ions throughout the gradient. Unscheduled adductomic method: all targeted nucleosides were monitored throughout the gradient. Scheduled adductomic method: targeted nucleosides were monitored during a 6 min window around their predicted RT. All measurements were conducted using a PGC trapping column and a Zorbax SB-C18 analytical column. Peak area ratios were normalized to the average peak area ratio obtained with the cPu method. Mean and standard deviation are plotted for each group ($n = 4$). There were no statistically significant differences in the normalized peak area ratios obtained from different methods.