HHS Public Access

Author manuscript

Anal Chem. Author manuscript; available in PMC 2019 December 04.

Published in final edited form as:

Anal Chem. 2019 December 03; 91(23): 15193–15203. doi:10.1021/acs.analchem.9b04068.

DNA Crosslinkomics: A Tool for the Comprehensive Assessment of Interstrand Crosslinks Using High Resolution Mass Spectrometry

Chiung-Wen Hu[†,#], **Yuan-Jhe Chang**[‡,#], **Marcus S. Cooke**[§,‖], **Mu-Rong Chao**[*,‡,⊥]

[†]Department of Public Health, Chung Shan Medical University, Taichung 402, Taiwan

[‡]Department of Occupational Safety and Health, Chung Shan Medical University, Taichung 402, Taiwan

[§]Oxidative Stress Group, Department of Environmental Health Sciences, Florida International University, Miami, Florida 33199, United States

[‖]Biomolecular Sciences Institute, Florida International University, Miami, Florida 33199, United States

[⊥]Department of Occupational Medicine, Chung Shan Medical University Hospital, Taichung 402, Taiwan

Abstract

DNA–DNA crosslinks, especially interstrand crosslinks (ICLs), cause cytotoxicity via blocking replication and transcription. Most measurements of ICLs lack sensitivity and structural information. Here, a high resolution, accurate mass spectrometry (HRMS) method was developed to comprehensively determine the untargeted, totality of DNA crosslinks, a.k.a. DNA crosslinkomics. Two novel features were introduced into this method: the accurate mass neutral losses of both two 2-deoxyribose (dR) and one dR groups will screen for ICLs as modified dinucleosides; the accurate mass neutral losses of both of the two nucleobases and one nucleobase will detect unstable DNA crosslinks, that could undergo depurination. Our crosslinkomics approach was tested by screening for crosslinks in formaldehyde- and chlorambucil-treated calf thymus DNA. The results showed that all expected drug-bridged crosslinks were detected successfully, along with various unexpected crosslinks. Using HRMS, the molecular formula and

[*]**Corresponding Author:** Fax: +886-4-23248194. chaomurong@gmail.com, mrchao@csmu.edu.tw. Corresponding author address: Department of Occupational Safety and Health, Chung Shan Medical University, Taichung 402, Taiwan.

[#]Author Contributions

C.-W.H. and Y.-J.C. contributed equally to this work.

The authors declare no competing financial interest.

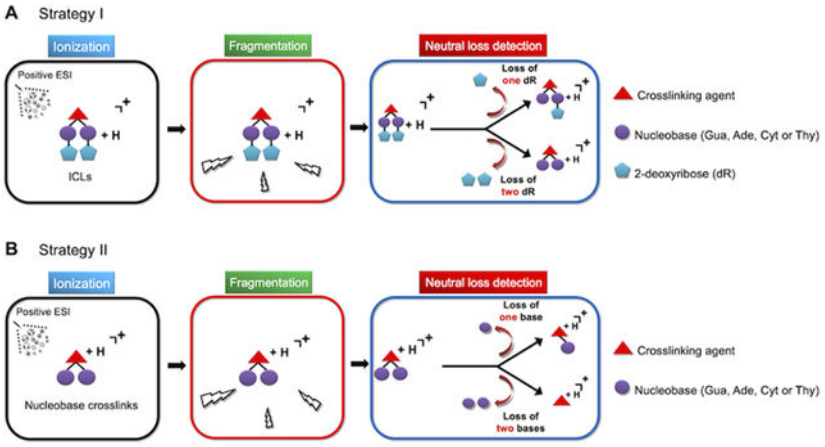
ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.9b04068. Experimental details for CT-DNA treated with formaldehyde and CLB; dose–response relationships in treatment of formaldehyde or CLB; product ion spectra of CLB- or formaldehyde-induced DNA ICLs; comparisons of different enzyme combinations of hydrolysis in releasing DNA crosslinks; comparison of acid-thermal hydrolysis and enzymatic hydrolysis in releasing DNA crosslinks; and confirmation ions for each DNA crosslink detected induced by formaldehyde and CLB (PDF)

chemical structures of these unexpected crosslinks were determined. The formation of apurinic/apyrimidinic (AP) site-derived crosslinks, at levels comparable to those for drug-bridged crosslinks, highlighted their novel, potential role in cytotoxicity. Our new crosslinkomics approach can detect expected and unexpected environmental and drug-induced crosslinks in biological samples. This broadens the existing cellular DNA adductome and offers the potential to become a powerful tool in precision medicine.

Graphical Abstract



The genome is constantly exposed to various endogenous and exogenous sources of DNA-modifying agents, some of which have two reactive groups (i.e., bifunctional agents). In addition to generating monoadducts, these bifunctional agents may form a second bond with DNA, resulting in the formation of a crosslink.[1] There are two types of DNA crosslinks: those that are within the same DNA strand (intrastrand crosslinks) and those that are between the opposing stands of the DNA helix (interstrand crosslinks, ICLs). DNA ICLs are among the most toxic of DNA lesions since they prevent separation of the two DNA strands, which is a prerequisite for transcription and replication. Hence, ICLs act as an absolute block to essential cellular processes and are particularly detrimental to rapidly dividing cells. This has led to the extensive use of crosslinking agents, such as nitrogen mustards, cisplatin, mitomycin C, and psoralens, as potent anticancer therapies.[2]

In addition to the ICL-forming therapeutic drugs, other exogenous sources of ICLs include environmental pollutants, such as 1,3-butadiene from fuel combustion.[3] There are also endogenous sources of ICLs derived from the byproducts of lipid peroxidation, including malondialdehyde and crotonaldehyde, the concentrations of which are also strongly influenced by exogenous agents, such as the ingestion of dietary lipids and alcohol, and exposure to tobacco smoke and automobile exhaust.[4,5] Another agent of endogenous origin that has been shown to introduce ICLs is nitric oxide, a signaling molecule important for vasoregulation.[6]

In addition to the drug-bridged ICLs, a new class of ICLs has been recently identified in duplex DNA from the reaction of an apurinic/apyrimidinic (AP) site, also known as an abasic site, with nucleobases on the opposing strand.[7] AP sites are common DNA lesions

caused by spontaneous, or enzyme-catalyzed, hydrolysis of the *N*-glycosidic bond in DNA, and they exist in an equilibrium between a ring-open aldehyde (at the C1′ position) and a ring-closed hemiacetal.[8] The ring-open aldehyde is able to form ICLs by reacting with the exocyclic amino group of adenine (Ade) or guanine (Gua) residues on the opposite strand.[8,9] AP sites could be also derived from the oxidation of 2-deoxyribose moiety (e.g., C4′- or C5′-oxidized AP site) and have been shown to generate ICLs in duplex DNA.[10,11]

Mass spectrometry (MS) is the most sensitive and specific method for detecting and quantifying DNA lesions, and MS is also an invaluable tool for structurally identifying new adducts formed by reactive genotoxins.[12] DNA ICLs are challenging to detect and quantify in biological samples for a variety of reasons. In addition to their relative rarity, compared to other types of lesions, many ICLs (e.g., formed by the alkylation at the N7 position of Gua and the N7/N3 position of Ade) induced by alkylating agents (e.g., nitrogen mustards and nitrosoureas) are hydrolytically unstable, leading to spontaneous depurination. Spontaneous hydrolysis can take place during cell storage, DNA extraction, and enzymatic hydrolysis and leads to the loss of crosslinked nucleobases in the ICLs.[13] Although MS has been previously applied to measure chemotherapy-induced ICLs[14] and those derived from environmental pollutants,[3] it has not yet been applied to detect DNA crosslinks generated from endogenous cellular processes, not least due to the elusive identity of these crosslinks. Moreover, because of technical constraints, most MS-based methods for ICL measurement utilize targeted approaches, aiming at the determination of a single or a few specific ICLs formed from specific exposures, such as cancer chemotherapeutic agents. To the best of our knowledge, there has never been a comprehensive, untargeted analysis of DNA ICLs reported in the literature, akin to DNA adductomics.[15]

In the present study, we developed a high resolution, accurate mass spectrometry approach to comprehensively measure DNA ICLs (i.e., both expected and unexpected) in biological samples, and we coined the term “DNA crosslinkomics” to describe this approach. Our strategy for performing crosslinkomics is illustrated in Figure 1. The key step in this assay is based on the understanding that the DNA hydrolysis enzymes used can release the ICLs and intrastrand crosslinks as modified dinucleosides and dinucleoside phosphates, respectively.[16,17] Given that ICLs consist of two 2-deoxyribose (dR) moieties, and the 2′-deoxynucleoside glycosidic bond is extremely labile during positive ionization,[18–20] a key, novel feature of our DNA crosslinkomics approach is to monitor the accurate mass neutral loss of both one dR group (116.0474 amu) and two dR groups (232.0948 amu) (see Supporting Information, Figure S1A, Strategy I). Since the modified dinucleoside phosphates derived from intrastrand crosslinks are not readily protonated in positive ion mode,[21,22] they will not interfere with the measurement of the ICLs using Strategy I. Moreover, some purine nucleosides, modified at certain nucleobase positions (e.g., crosslinking between two Gua-N7 positions),[3] are unstable and may undergo spontaneous depurination during sample preparation, resulting in the formation of dinucleobase adducts (nucleobase crosslinks, Figure 1). To address this, another novel feature of our DNA crosslinkomics approach is the detection of nucleobase crosslinks by monitoring the neutral losses of one nucleobase as well as of any two nucleobases (e.g., 302.0988 amu for the loss of two Gua bases and 286.1038 amu for the loss of one Gua and one Ade nucleobases), in which the crosslinking modifications are positively charged (Figure S1B, Strategy II).

Because both ICLs and intrastrand crosslinks are comprised of dinucleobases (expect AP site-associated DNA crosslinks), Strategy II is unable to discriminate between interand intrastrand crosslinks.

The present DNA crosslinkomic approach was successfully tested by screening for DNA crosslinks induced by formaldehyde and chlorambucil (CLB, an alkylating chemotherapy drug) in calf thymus DNA (CT-DNA). For the formaldehyde treatment, not only were the five well-known DNA ICLs detected but also two unexpected DNA ICLs. In the case of CLB, two known DNA ICLs were detected, together with eight unexpected, novel DNA ICLs, illustrating the strength of the crosslinkomics approach.

EXPERIMENTAL SECTION

Chemicals

Methanol (MeOH), ethanol, 2-propanol, and tris-hydrochloride (Tris-HCl) were from either Macron Fine Chemicals or J.T. Baker (Center Valley, PA, USA). Ammonium acetate (AA), formic acid (FA), dimethyl sulfoxide (DMSO), sodium iodide (NaI), deferoxamine mesylate (DFO), nuclease P1, snake venom phosphodiesterase I (*Crotalus atrox*), formaldehyde, CLB, CT-DNA, and 2′-deoxyguanosine (dG) were from Sigma-Aldrich (St. Louis, MO, USA). Alkaline phosphatase and dephosphorylation buffer were from Roche (Grenzacherstrasse, Basel, Switzerland). [$^{15}N_5$]-8-Oxo-7,8-dihydro-2′-deoxyguanosine ([$^{15}N_5$]-8-oxodG) was from Cambridge Isotope Laboratories (Tewksbury, MA, USA). [d_4]-N^6-(2-Hydroxyethyl)-2′-deoxyadenosine ([d_4]-N^6-2HEdA) was from Toronto Research Chemicals (Ontario, Canada). N^6-Benzoyl-8-hydroxy-2′-deoxyadenosine (N^6-Bz-8-oxodA) was from Carbosynth (Compton, UK).

Formations of DNA Crosslinks Induced by Formaldehyde or CLB in CT-DNA

Reactions of CT-DNA with formaldehyde[23] or CLB[24] were performed as previously described. Full details are provided in the Supporting Information.

DNA Purification and Enzymatic Hydrolysis

After incubation with formaldehyde or CLB, DNA was precipitated and recovered using a solution of NaI, 2-propanol, and ethanol, as previously described.[25] The DNA pellet obtained was dissolved in 500 μL of 0.1 mM DFO solution, to a DNA concentration of ~500 μg/mL. The DNA recovery was nearly 100%, as determined by the dG content using LC-MS/MS.[26]

For DNA enzymatic hydrolysis, a previously reported three enzyme combination protocol[16] (nuclease P1, alkaline phosphatase and snake venom phosphodiesterase I) was employed to selectively release the ICLs (as modified dinucleosides) and intrastrand crosslinks (as modified dinucleoside phosphates). Full details are given in the Supporting Information.

Liquid Chromatography-High Resolution Mass Spectrometry

The DNA crosslinks were separated by a reversedphase HPLC system (Waters, Acquity UPLC, Milford, MA, USA) using an Inertsil ODS-3 C18 column (150 × 2.1 mm i.d., 5 μm,

GL Sciences, Tokyo, Japan). Full details regarding LC conditions are given in the Supporting Information. Mass spectrometry analysis was performed on an LTQ Orbitrap Fourier transform mass spectrometry (FTMS) instrument (LTQ Orbitrap Elite, Thermo Fisher Scientific, MA, USA) equipped with an HESI-II electrospray source, operated in positive ion mode. The source voltage was 3.5 kV. Both the heater and capillary temperatures were 200 °C. The gas setting of sheath gas, auxiliary gas, and sweep gas were 30, 15, and 1 arbitrary unit, respectively. The S-lens level was set at 40%. Data acquisition and processing were conducted by Xcalibur software 2.2 (Thermo Fisher Scientific).

Data-Dependent Neutral Loss-Driven MS^3 (DD-NL-MS^3) Acquisition and Analysis

Strategy I: Screening for ICLs as Modified Dinucleosides—The DD-NL-MS^3 scanning was initiated with repeated full scan detection in the FTMS, followed by MS^2 acquisition and constant neutral loss triggering of MS^3 fragmentation. The FTMS full scan ranged from *m*/*z* 220 to 800 and was performed in the Orbitrap at a resolution of 60,000 with automatic gain control (AGC) of 1×10^6 and a maximum ion injection time of 150 ms. The three most intense full scan ions together with the ion intensity greater than 7,500 from the MS^1 full scan spectra were selected and fragmented in the ion trap (MS^2 fragmentation by collision-induced dissociation, CID), at a resolution of 15,000. The normal 2′-deoxyribonucleosides and their ion adducts (e.g., $[M + Na]^+$, $[M + K]^+$, $[M + NH_3]^+$, and $[M + H + H_2O]^+$) were excluded from the MS^2 fragmentation. MS^2 fragmentation parameters were as follows: the ion injection time was 50 ms, the isolation mass width was 2 amu, the normalized collision energy (NCE) was set at 40%, and the AGC was 5×10^4. Data-dependent acquisition parameters were as follows: dynamic exclusion repeat count of one, repeat duration of 30 s, exclusion list size of 250, exclusion time of 9 s, and exclusion mass width of ±5 ppm. MS^3 fragmentation utilizing higher collision energy dissociation (HCD), using a NCE of 50%, was triggered if a neutral loss of two dR moieties (232.0948 amu ± 5 ppm, see Strategy I in Figure S1A) from the parent ion was detected, and one of the 10 most intense product ions from the MS^2, giving a minimal signal of 250, was observed. Except for NCE, other MS^3 fragmentation parameters (i.e., resolution, ion injection time, isolation mass width, and AGC value) were the same as those used for the MS^2 fragmentation event, as described above. All spectra were acquired using two common background ion signals, including *m*/*z* at 391.284290 [bis(2-ethylhexyl) phthalate] and *m*/*z* at 445.120030 (dodecamethylcyclohexasiloxane) as lock masses to ensure mass accuracy.

Strategy II: Screening for Nucleobase Crosslinks as Modified Dinucleobases—The FTMS full scan ranged from *m*/*z* 220 to 800 and was conducted at a resolution of 60,000, followed by MS^2 CID fragmentation at NCE of 40% with wideband activation. The parameters and exclusion criteria of MS^2 fragmentation and data-dependent acquisition were the same as for Strategy I, as described above. The subsequent MS^3 HCD fragmentation (NCE of 70%) was triggered if in the MS^2 a neutral loss of two nucleobases from the parent ion was detected within a mass tolerance of 5 ppm (see Strategy II in Figure S1B), including 302.0988 amu (Gua/Gua), 270.1088 amu (Ade/Ade), 222.0864 amu (Cyt/Cyt), 252.0858 amu (Thy/Thy), 286.1038 amu (Gua/Ade), 262.0926 amu (Gua/Cyt), 277.0923 amu (Gua/Thy), 246.0976 amu (Ade/Cyt), 261.0973 amu (Ade/Thy), and 237.0861 amu (Cyt/Thy). MS^3 HCD fragmentation parameters were the same as those applied for Strategy I.

Data Processing, Peak Alignment, and Normalization

Raw data files generated from the DD-NL-MS^3 scan were directly processed using the MZmine 2. Initially, the MS^3 peak list was generated showing all the precursors that had triggered the MS^3 scan event, owing to the detection of NL of two dR (or two nucleobases) in the MS^2 fragmentation spectrum, using the MS/MS peak list builder module. MS^3 spectra were extracted from the raw data, and the *m*/*z* and retention time (RT) tolerance were 0.001 amu and 0.5 min, respectively. Then the chromatographic data for the MS^3 peak list was reconstructed by using the peak extender module with a *m*/*z* tolerance of 5 ppm and minimum height of 250. Chromatograms were deisotoped using the isotopic peaks grouper algorithm, with a *m*/*z* tolerance of 5 ppm and a RT tolerance of 0.75 min. In addition, duplicate peaks were filtered using the duplicate peak filter module, with a *m*/*z* tolerance of 5 ppm and a RT tolerance of 0.75 min. The RT normalizer was adopted to decrease deviation of RTs between peak lists with the parameter setting as follows: *m*/*z* tolerance of 5 ppm, RT tolerance of 0.75 min, and the minimum peak intensity of 5×10^5. Peak alignment (including precursor *m*/*z* and RT) in different samples was performed using the Join aligner module, with a *m*/*z* tolerance of 5 ppm and RT tolerance of 0.5 min. Finally, a peak list report was created showing all aligned/normalized peaks *m*/*z*, RT, and intensity of the precursor ions that had triggered DD-NL-MS^3 fragmentation. The resulting peak list, saved as a CSV file, was then exported to Microsoft Excel. The peak intensity was then normalized to correct for variations due to matrix effects (ME), by adjusting the peak intensity in full scan of analyte with those of internal standards spiked, including [$^{15}N_5$]-8-oxodG (*m*/*z* 289.0841, for correcting ME during RT 0–20 min), [d_4]-N^6-2HEdA (*m*/*z* 300.1604, for correcting ME during RT 20–35 min), and N^6-Bz-8-oxodA (*m*/*z* 372.1302, for correcting ME during RT 35–60 min), as previously described by Chang et al.[27] The aligned/normalized peaks were converted to a visual, colored DNA crosslinkome map, which consisted of RT (X axis), *m*/*z* (Y axis), and adjusted peak area (spot color), using OriginPro 2016 Sr2 software (MA, USA). The results were presented as mean values of duplicates of two independent experiments.

Confirmation of the Candidate DNA Crosslinks

All the ions that had triggered an MS^3 fragmentation event were considered as candidate DNA crosslinks and were further confirmed using their specific fragments observed in the MS^2 and MS^3 spectra. For ICLs measured by Strategy I, in addition to the fragment that lost two dR detected in the MS^2 spectra, the candidates must meet the criteria of observations of the fragment that lost one dR in the MS^2 spectra as well as the fragment that corresponded to the nucleobase moiety in the MS^3 spectra. For nucleobase crosslinks, measured by Strategy II, in addition to the fragment losing two nucleobases detected in the MS^2 spectra, the candidates must have the fragment losing one nucleobase in the MS^2 spectra, as well as the fragment corresponding to the crosslinking agent moiety in the MS^3 spectra.

RESULTS

Screening of Formaldehyde-Induced Expected and Unexpected DNA Crosslinks Using LC-HRMS in DD-NL-MS^3 Scan Mode

After formaldehyde treatment, the DNA samples were subjected to enzymatic hydrolysis and analyzed using Strategies I and II as described above. Using Strategy I, MS^3 fragmentation was triggered if a neutral loss of two dR moieties (232.0948 amu) from the parent ion was detected in the MS^2. In comparison to the control sample (Figure 2A), a total of eight specific ions were observed on the DNA crosslinkome map in the formaldehyde-treated DNA samples (Figure 2B). The ions detected are summarized in Supporting Information Table S1. Of the eight ions, two were symmetrical ICLs involving two dG or two dA, four were asymmetrical ICLs involving dG, dA, and dC, and two ICLs were formed between dA and an AP site. The chemical structures of these formaldehyde induced ICLs are shown in Figure 3.

The two symmetrical ICLs (Ions 1 and 4) yielded their protonated precursor ions $[M + H]^+$ of two dG or dA, each, with the addition of 14.0156 amu (a CH_2 group, from the formaldehyde). For example, Ion 1 (m/z 547.2011) eluted at 28.17 min and was referred to as the ICL dG-CH_2-dG, formed by formaldehyde. As shown in Figure 4A, Ion 1 was confirmed by the observation of specific fragment ions of m/z 315.1058 ($[M + H - 2dR]^+$), m/z 431.1548 ($[M + H - dR]^+$), and m/z 268.1038 ($[dG + H]^+$) in the MS^2, as well as the specific fragment ion of m/z 152.0566 ($[Gua + H]^+$) in the MS^3 spectrum. The four asymmetrical crosslinks (ions 2, 3, 5, and 6) involving (dG, dA, and dC) were proposed to be dG-CH_2-dA, dA-CH_2-hydroxymethyl-2′-deoxycytidine (hmdC), dG-CH_2-dC, and dA-CH_2-dC, respectively. For example (Figure 4B), for Ion 6 (dA-CH_2-dC, which eluted at 27.8 min), a dominant peak corresponding to $[M + H]^+$ of 491.1999 was detected in full scan. This ion triggered a MS^2 fragmentation event, and the observation of ions corresponding to the neutral loss of two dR (m/z 259.1048) and one dR (m/z 375.1521) in the MS^2, as well as the ions corresponding to $[Ade + H]^+$ (m/z 136.0620) and $[Cyt + H]^+$ (m/z 112.0504) in the MS^3 spectrum, confirmed the identity of dA-CH_2-dC. Furthermore, Ions 7a and 7b were isomers and were identified to be dA-AP crosslinks. This assumption was supported by the presence of a protonated precursor ion $[M + H]^+$ of 368.1563 (Figure 4C), the specific ions of m/z 136.0615 and 252.1086 in the MS^2 due to the neutral loss of two dR and one dR, and a further fragment ion $[Ade + H]^+$ of m/z 136.0615 in the MS^3. Ion 7a was also detected in the control DNA sample but at a lower intensity compared to that observed in the formaldehyde-treated DNA. The rank order of formation (in terms of signal intensity) of seven crosslinks was as follows: dA-CH_2-dA > dA-CH_2-dC > dG-CH_2-dA > dG-CH_2-dG > dG-CH_2-dC > dA-CH_2-hmdC or dA-AP, as shown in Table S1. The formation of these crosslinks increased in a dose-responsive manner (Supporting Information, Figure S2).

When Strategy II was applied, no ions were detected in either the formaldehyde-treated (Figure 2E) or the control samples (see Figure 2D). We speculate that this is probably because no ions lost two nucleobases and therefore did not trigger the MS^3 fragmentation. This may further indicate that formaldehyde-induced crosslinks are stable and do not undergo depurination during enzymatic hydrolysis.

Screening of CLB-Induced Expected and Unexpected DNA Crosslinks Using LC-HRMS in DD-NL-MS^3 Scan Mode

DNA crosslinkome maps obtained from CLB-treated DNA samples, using Strategy I, are shown in Figure 2C. A total of 19 ions (Ions 8–17) were newly formed following the CLB treatment, and the presence and intensity of the diagnostic precursor and fragment ions of each crosslink are listed in Supporting Information Table S2. The chemical structures of CLB-induced ICLs, measured as modified dinucleosides, are shown in Figure 5. Four ions were proposed to be symmetrical ICLs, formed by CLB, including dG-CLB-dG (Ions 8a and 8b), dA-CLB-dA (Ion 10), and dC-CLB-dC (Ion 13). For example, Ion 8a (eluted at 19.35 min) and Ion 8b (at 34.74 min) were isomers, having the same protonated precursor ion $[M + H]^+$ of 766.3270 and the same fragmentation pattern (Figure S3A); both isomers had the characteristic ions of m/z 534.2318 ($[M + H - 2dR]^+$) and m/z 650.2787 ($[M + H - dR]^+$) yielded in the MS^2 and the consecutive fragment ion of m/z 178.0725 ($[Gua + C_2H_2]^+$) coming from the Gua moiety and the CLB fragment ion at m/z 232.1325 ($[M + H - 2dG]^+$) due to the elimination of two dG in the MS^3 spectrum, supporting the presence of the dG-CLB-dG crosslink. Moreover, Ions 9, 11, and 12 represented asymmetrical ICLs, which were dG-CLB-dA, dG-CLB-dC, and dA-CLB-dC, respectively. In the case of CLB bridged from Gua to Ade (Ions 9a and 9b, Figure S3B), our identification of the ICL was based on the detection of the precursor ion $[M + H]^+$ at m/z 750.3324 in the full scan, and the following characteristic ions of m/z 518.2372 ($[M + H - 2dR]^+$) and m/z 634.2840 ($[M + H - dR]^+$) in the MS^2, and the appearance of fragment ions of m/z 178.0721 ($[Gua + C_2H_2]^+$), m/z 162.0772 ($[Ade + C_2H_2]^+$), and m/z 232.1337 ($[M + H - dG - dA]^+$), corresponding to the Gua moiety, Ade moiety, or CLB group, respectively, in the MS^3 spectrum.

Meanwhile, a novel ICL (Ions 14a and 14b) was newly identified, which we propose is dG-CLB-dR. This conclusion was based on the appearance of $[M + H]^+$ at 633.2877, together with the specific ions of m/z 401.1932 ($[M + H - 2dR]^+$) and m/z 517.2401 ($[M + H - dR]^+$) in the MS^2, and the following fragment ion of m/z 250.1433 ($[M + H - dR - dG]^+$) corresponding to the CLB moiety in the MS^3 spectrum (see Figure S3C). Furthermore, Ions 15, 16, and 17 were also identified to be ICLs formed by the reaction of an AP site with dA, 5-methyl-2′-deoxycytidine (5-medC), or dC, respectively. Figure S4A illustrates an example of the MS/MS spectrum obtained, confirming the identification of the dC-AP crosslink (Ion 17), while the detailed MS/MS fragmentation of 5-medC-AP (Ion 16) is provided in Figure S4B.

It was interesting to note that the dA-AP crosslinks induced by CLB had four isomers (Ions 15a, 15b, 15c, and 15d) and that Ion 15c and Ion 15d were also detected in the formaldehyde-treated DNA (where they were referred to as Ion 7a and Ion 7b). The rank order of formation of CLBinduced ICLs (measured as dinucleoside adducts using Strategy I) was as follows: dA-AP > dC-AP > dG-CLB-dG > dG-CLB-dA or dG-CLB-dC > dG-CLB-dR > 5-medC-AP > dA-CLB-dC or dA-CLB-dA or dC-CLB-dC (Table S2). Furthermore, most of these crosslinks increased in a dose-responsive manner (see Figure S5).

CLB-treated DNA samples were also analyzed by Strategy II (Figure S1B), in which the MS^3 fragmentation event was triggered by the observation of neutral loss of any two nucleobases during the MS^2 fragmentation. The resulting DNA crosslinkome map is shown

in Figure 2F. A total of seven ions (as modified dinucleobases, Ions 18–22) were newly identified and listed in Supporting Information Table S3. The chemical structures of CLB-induced crosslinks measured as modified dinucleobases are shown in Figure S6. Two symmetrical crosslinks were characterized to be Gua-CLB-Gua (Ions 18a, 18b, and 18c) and Ade-CLB-Ade (Ion 20), while the remaining ones were asymmetrical crosslinks, including GuaCLB-Ade (Ion 19), Gua-CLB-Cyt (Ion 21), and Ade-CLB-Cyt (Ion 22). As shown in Figure 6, a general feature in the MS^2 fragmentation of the modified dinucleobases was the neutral loss of one and two of the crosslinked nucleobases, yielding the modified nucleobase $[M + H - B]^+$ and the remaining CLB moiety $[M + H - 2B]^+$, respectively, while in the MS^3 fragmentation, it gave rise to the various fragments of CLB moiety. The rank order of formation of CLB-induced crosslinks (measured as modified dinucleobases using Strategy II) was as follows: Gua-CLB-Ade > Gua-CLB-Gua > Ade-CLB-Ade > Ade-CLB-Cyt, or Gua-CLB-Cyt (Table S3).

DISCUSSION

For decades, DNA crosslinks were measured mostly by biochemical and biophysical techniques, including the alkaline comet assay, the agarose gel-based method, and the alkaline filter elution.[28,29] However, these methods do not allow for the absolute quantification of ICLs, often lack sensitivity, and do not provide structural information about the nature of the ICLs. It is also reported that considerable differences in the assay protocols used by different research groups affect the results when interlaboratory comparisons are made.[30] With the rapid development of mass spectrometry techniques, we herein proposed a novel crosslinkomics approach using LC-HRMS to screen for the totality of DNA ICLs.

The present approach relied on the assumption that the enzymes used for DNA hydrolysis can release the ICLs and intrastrand crosslinks as modified dinucleosides and dinucleoside phosphates, respectively. Our approach used a combination of three enzymes in the DNA hydrolysis protocol (nuclease P1, alkaline phosphatase, and snake venom phosphodiesterase I), which has been shown to selectively release ICLs and intrastrand crosslinks induced by mitomycin C in CT-DNA.[16] However, we also compared two additional protocols for DNA hydrolysis, including a two enzyme combination (nuclease P1 and alkaline phosphatase)[26] and a four enzyme combination (nuclease P1, alkaline phosphatase, snake venom phosphodiesterase I, and bovine spleen phosphodiesterase II).[9] The results are shown in the Supporting Information (Figures S7 and S8). In the case of formaldehyde induced ICLs (Figure S7), the results clearly showed that when DNA was digested using a two enzyme combination protocol, dG-CH_2-dG was not released, and the other ICLs were detected in significantly lower amounts than when three enzyme and four enzyme combination protocols were used. Similar findings were also found in the CLB-treated DNA (Figure S8), showing that lower amounts of ICLs were detected, and six ICLs were not even released using the two enzyme combination protocol. Our observations are in line with previous findings showing that nuclease P1 alone is not capable of efficiently cleaving the 3′-phosphodiester bond of ICLs formed by mitomycin C or psoralen derivatives.[12,31] We noted no significant difference in both the amounts and forms of ICLs released using the three enzyme and four enzyme combination protocols. Nevertheless, since no synthetic DNA oligomer standards containing ICLs and intrastrand crosslinks were used in this study, it is

difficult to ensure that 100% of the induced ICLs were released using the three or four enzyme combination protocols. There could also be a possibility that the enzymes used may release the intrastrand crosslinks by excising the phosphate group (yielding the same digested product as those of ICLs)[32] and therefore may have contributed to the levels of ICLs detected. Subsequently, without removing or capturing the monoadducts prior to enzymatic hydrolysis, it is also possible that the ICLs detected might result from the reaction of a monoadduct with another unmodified mononucleoside in the nucleoside mixture during and/or after enzymatic hydrolysis. We acknowledge the two above potential limitations of the present study.

Intrastrand crosslinks could be released as modified dinucleoside phosphates using the three enzyme combination protocol.[16] As expected, no modified dinucleoside phosphates were able to trigger MS^3 fragmentation. This is probably because the modified dinucleoside phosphates, derived from intrastrand crosslinks, are not readily protonated in positive ion mode.[21,22] Furthermore, even if the modified dinucleoside phosphates were protonated in positive ion mode, they would have a different fragmentation pattern compared to ICLs. A previous study showed that the modified dinucleoside phosphates tend to lose one dR group (116.0474 amu) and two dR with a phosphate group (294.0505 amu) under positive ionization.[33] This means that the modified dinucleoside phosphates will not undergo neutral loss of two dR groups upon fragmentation in the positive ionization, which is the trigger for MS^3 fragmentation (Strategy I). On this basis, they will not interfere with the measurement of ICLs by our LC-DD-NL-MS^3-based DNA crosslinkomics assay.

Using our crosslinkomics approach, we investigated the formation of formaldehyde- and CLB-induced DNA ICLs. In the case of formaldehyde, two symmetrical ICLs (dG-CH_2-dG and dA-CH_2-dA) and three asymmetrical ICLs (dG-CH_2-dA, dG-CH_2-dC, and dA-CH_2-dC) previously reported in the literature were successfully identified in this study.[34] Among these ICLs, dA-CH_2-dA (Ion 4) was formed in the highest yield, again similar to previous findings.[23] Our method also detected two unexpected ICLs formed by formaldehyde, dA-CH_2-hmdC and dA-AP. Detailed MS^2 and MS^3 spectra of dA-CH_2-hmdC (Ion 3) are given in Figure S9. Based on the appearance of the $[hmdC + H]^+$ fragment (m/z 258.1080), we speculate that dA-CH_2-hmdC is formed via the initial generation of the well-known N^4-hydroxymethyl-dC, followed by the formation of a methylene-bridge to dA.

In addition to the methylene-bridged crosslinks, this study also observed the dA-AP crosslink in the formaldehyde-treated DNA (Figure 4C). This class of crosslinks, comprising a 2′-deoxyribonucleoside linked to an apurinic/apyrimidinic site, includes dA-AP and dG-AP and was first identified by the Gates research group[8,9,35] in duplex DNA using NMR and low resolution MS. It has been suggested that this class of ICLs has persistent lesions (e.g., half-life: 3–4 days for dA-AP and 22 days for dG-AP at neutral pH) with the potential to block activities associated with DNA processing (e.g., replication). It is generally believed that the level of DNA ICLs in a normal cell is very low[36] and rarely, or barely, detectable. However, the present study showed that control, untreated CT-DNA had a detectable amount of dA-AP crosslinks, which led us to speculate that these dA-AP may be formed during the manufacturing process (e.g., DNA isolation). Artifactual formation of AP sites during DNA isolation could derive from the BER enzyme activity, spontaneous depurination at high

temperature and/or low pH, and β-elimination.[37] These artifactually derived AP sites might result in the formation of further AP site-associated DNA crosslinks.

It has been reported that reactive oxygen species (ROS) can be formed during DNA extraction, depending upon the method used, and lead to the artefactual formation of oxidatively generated damage to DNA.[38] Although hydrolysis of the *N*-glycosidic bond of normal purine nucleosides is likely to contribute to the formation of AP sites, ROS have also been shown to contribute to the formation of AP sites.[37,39] To test our hypothesis that the dA-AP observed in CT-DNA arose mainly during the DNA extraction process, we investigated the formation of DNA crosslinks in liver tissue obtained from control mouse and rat. The liver DNA was extracted using the protocol suggested by ESCODD[40] to minimize/eliminate artifactual oxidation of DNA. Our results show that no DNA crosslinks were detectable in the liver tissue DNA from both the control mouse or rat (data not shown, i.e., the appropriate spot was absent from the crosslinkome map). We also measured the cellular level of 8-oxo-7,8-dihydro-2′-deoxyguanosine (8-oxodG, a biomarker of oxidatively damaged DNA) in CT-DNA and liver tissues of mouse/rat using our previously reported LC-MS/MS method.[26] The results showed that the 8-oxodG level in liver DNA was as low as <5 8-oxodG/10^6 dG (or 3.7 8-oxodG/10^6 dG in untreated cell[40]), whereas in CT-DNA it was 150 8-oxodG/10^6 dG. Combined, these findings illustrate that not only do ROS induce “classical” forms of oxidatively generated DNA damage, such as 8-oxodG, but they may also generate AP site-associated DNA crosslinks. This finding also warns of taking suitable precautions to avoid artifact formation during sample workup for DNA crosslink analysis (including for the gel-based assays).

Nitrogen mustard CLB is an anticancer drug and has been used clinically against chronic lymphatic leukemia, lymphomas, and advanced ovarian and breast carcinomas.[41] The cytotoxicity of CLB is understood to be due to the alkylation and crosslinking of DNA. CLB, being an alkylating agent, preferentially binds to the DNA at specific locations, such as the N7 of guanine and the N3 of adenine. Alkylation at these endocyclic nitrogens of the nucleobase can destabilize the nucleobases and cause depurination, resulting in the AP site formation.[42] Additionally, CLB possesses two electrophilic sites and can form a crosslink between two nucleophilic centers in the DNA duplex, representing the most toxic of all alkylation events.

In this study, we successfully detected the expected CLB-linked ICLs, including dG-CLB-dG and dG-CLB-dC, as previously reported.[24,43] Excitingly, we identified five unexpected CLB-linked ICLs which have not been reported in the literature, (i.e., dG-CLB-dA, dA-CLB-dA, dA-CLB-dC, dC-CLB-dC, and dG-CLB-dR, see Figure 5). In addition to the formation of drug-bridged ICLs, our method also revealed the formation of AP-derived ICLs, including dA-AP, dC-AP, and 5-medC-AP. Of these three AP-derived ICLs, dC-AP and 5-medC-AP were identified here for the first time. Our findings answer a question raised by the Gates laboratory as to whether the AP sites opposed by a cytosine residue could engage in the interstrand crosslink reaction,[35] and clearly they can. We further suggest that this reaction not only occurs with cytosine residues but also with the 5-medC (the so-called fifth nucleoside, which plays a major role as an epigenetic modification of DNA). Furthermore, it is interesting to note that the amounts of AP-derived ICLs generated are

relatively higher than drug-bridged crosslinks (in terms of peak intensity). This finding implies that, in addition to the drug-bridged ICLs, the AP-derived ICLs could also play a critical role in the cytotoxicity of DNA alkylating drug and warrants further investigation. Subsequently, it should be mentioned that, in this study, the AP-derived ICLs detected following formaldehyde or CLB treatments were derived from the C1′ of 2-deoxyribose moiety (ring-opened aldehyde group at the C1′ position) with another 2′-deoxynucleoside, which was supported by the observation of specific fragments that had lost two dR. Because C4′- or C5′-oxidized AP site-containing ICLs contain only one dR structure (not two dR), they could not be detected in the present study.

In the present study, we proposed two strategies (Figure S1): Strategy I is for the determination of ICLs (in the form of modified dinucleosides), while Strategy II is for the determination of nucleobase crosslinks. Strengths of Strategy I include the ability to determine th stable ICLs and the ability to distinguish between the stable ICLs and intrastrand crosslinks. However, a weakness is that the successful application of Strategy I relies entirely on the proper use of enzymatic hydrolysis to release the ICLs and intrastrand crosslinks as modified dinucleosides and dinucleoside phosphates, respectively. Strategy II is an approach to address the evaluation of those unstable DNA crosslinks which could undergo spontaneous depurination (to form nucleobase crosslinks) during sample preparation/storage. Alternatively, to avoid potential release bias of the hydrolysis enzymes, Strategy II can also be applied to DNA samples following acid and thermal hydrolysis, in which all the DNA crosslinks will be released as nucleobase crosslinks. An example of the use of Strategy II to measure the DNA crosslinks in the CLB-treated DNA following acid-thermal hydrolysis[25] is provided (Figure S10). Using Strategy II, the amounts of DNA crosslinks following acid-thermal hydrolysis were ~2–20-fold more than those following enzymatic hydrolysis. Furthermore, 16 more isomers of certain crosslinks were detected using acid-thermal hydrolysis as listed in Table S4, possibly resulting from the release of intrastrand crosslinks. Overall, compared to Strategy I with enzymatic hydrolysis, Strategy II using acid-thermal hydrolysis is relatively simple and fast for the determination of DNA crosslinks and therefore makes it more applicable to clinical practice. Nevertheless, the weaknesses of Strategy II are that the ICLs and intrastand crosslinks are not distinguishable, and the AP-site derived crosslinks cannot be detected, despite, as we have shown, this type of crosslink being present in considerable amounts. Furthermore, when Strategy II was applied to cellular DNA analysis (e.g., cells or tissues), without due care to extract DNA specifically, then some RNA crosslinks might have interfered with the measurement of DNA crosslinks. Previously, Stornetta et al.[13] developed an untargeted approach to screen for DNA alkylation mono- and crosslinked adducts. Their approach was to monitor the accurate mass neutral loss of one dR group or one of the four nucleobases, followed by triggering the MS^3 fragmentation. Although several drug-induced crosslinks were detected (i.e., drug bridging two purine residues), their approach was not specific for the measurement of DNA crosslinks, and clearly some important types of drug-induced crosslinks were overlooked, including the AP site-associated DNA crosslinks, drug bridging pyrimidine residues, and drug bridging one 2′-deoxyribonucleoside and one dR.

There is one type of alkylating drug-induced crosslink that has not been detected using either strategy following enzymatic hydrolysis, which comprises the drug bridging one 2′-

deoxynucleoside and one nucleobase, such as dG-CLB-Gua, dA-CLB-Gua, dG-CLB-Ade, dA-CLB-Ade, dC-CLB-Gua, and dC-CLB-Ade (as summarized in Table S5). We presume that this type of crosslink is formed when one modified nucleoside underwent depurination during/after enzymatic hydrolysis, while the other modified nucleoside did not. Although this type of crosslink is formed in similar amounts as ICLs (as modified dinucleosides, see Table S2), they are unable to trigger MS^3 fragmentation because they fail to meet the MS^3 fragmentation criteria of either Strategy I (loss of two dR) or Strategy II (loss of two nucleobases). However, this type of crosslinks shows a consistent pattern fragmentation (see Table S5 and Figure S11), which is the neutral loss of one dR in combination with two nucleobases. Such crosslinks could therefore be easily detected by setting the triggering criteria of MS^3 fragmentation accordingly, i.e., if in the MS^2 a neutral loss of two nucleobases with one dR from the parent ion was detected (e.g., 418.1462 amu for loss of one dR and two Gua and 402.1512 amu for loss of one dR, one Gua, and one Ade). This represents a potential Strategy III for future study.

CONCLUSION

We have developed a HRMS method for determination of the totality of DNA crosslinks, which we have termed DNA crosslinkomics. The proposed DNA crosslinkomics approach would be more easily performed when the mass spectrometry allows for setting more triggering criteria (e.g., Orbitrap Fusion Lumos Tribrid MS). For example, if all the triggering criteria (i.e., the neutral loss of two dR for Strategy I, neutral loss of two bases for Strategy II, and neutral loss of two bases/one dR for Strategy III) could be included in one run, all types of DNA crosslinks could be detected simultaneously following one injection, after enzymatic hydrolysis. A number of unexpected and novel DNA ICLs following formaldehyde and CLB treatments were discovered. However, their putative structures require further identification using independently synthesized standards, followed by MS and nuclear magnetic resonance spectroscopy for full structural characterization.

Little is known about the repair of crosslinks in human cells,[44] and it has long been asked whether the repair products of DNA crosslinks are further excreted into urine and, if yes, in what form would they be present in urine. As long as they retain their characteristic structures, then the repair products could conceivably be detected. As urine can be obtained noninvasively, unlike most tissue DNA, our DNA crosslinkomics approach, reported here, could also be applied to urine, providing a useful tool to help elucidate the molecular steps involved in the repair of ICLs, together with biomonitoring exposure to crosslinking agents. This method not only offers structural insights into ICLs, caused by a variety of endogenous metabolites and environmental exposures that have two reactive groups, but also provides a platform for the development of predictive diagnostic tests to improve the use of clinical crosslinking agents in precision medicine.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We greatly appreciate the assistance of Pei-Shan Li with experimental procedures.

Funding

This work was funded by the Ministry of Science and Technology, Taiwan [grant numbers MOST 106-2314-B-040-015-MY3 and NSC 102-2314-B-040-016-MY3]. The research reported in this publication was also supported, in part, by the National Institute of Environmental Health Sciences of the National Institutes of Health under award number R01ES030557.

REFERENCES

(1). Noll DM; Mason TM; Miller PS Chem. Rev. 2006, 106, 277–301. [PubMed: 16464006]

(2). Clauson C; Scharer OD; Niedernhofer L Cold Spring Harbor Perspect. Biol. 2013, 5, a012732.

(3). Sangaraju D; Goggin M; Walker V; Swenberg J; Tretyakova N Anal. Chem. 2012, 84, 1732–1739. [PubMed: 22220765]

(4). Deans AJ; West SC Nat. Rev. Cancer 2011, 11, 467–480. [PubMed: 21701511]

(5). Stone MP; Cho YJ; Huang H; Kim HY; Kozekov ID; Kozekova A; Wang H; Minko IG; Lloyd RS; Harris TM; et al. Acc. Chem. Res. 2008, 41, 793–804. [PubMed: 18500830]

(6). Caulfield JL; Wishnok JS; Tannenbaum SR Chem. Res. Toxicol. 2003, 16, 571–574. [PubMed: 12755585]

(7). Price NE; Li L; Gates KS; Wang YS Nucleic Acids Res. 2017, 45, 6486–6493. [PubMed: 28431012]

(8). Price NE; Catalano MJ; Liu S; Wang YS; Gates KS Nucleic Acids Res. 2015, 43, 3434–3441. [PubMed: 25779045]

(9). Catalano MJ; Liu S; Andersen N; Yang Z; Johnson KM; Price NE; Wang Y; Gates KS J. Am. Chem. Soc. 2015, 137, 3933–3945. [PubMed: 25710271]

(10). Regulus P; Duroux B; Bayle PA; Favier A; Cadet J; Ravanat JL Proc. Natl. Acad. Sci. U. S. A. 2007, 104, 14032–14037. [PubMed: 17715301]

(11). Guan L; Greenberg MM J. Am. Chem. Soc. 2009, 131, 15225–15231. [PubMed: 19807122]

(12). Liu S; Wang YS Chem. Soc. Rev. 2015, 44, 7829–7854. [PubMed: 26204249]

(13). Stornetta A; Villalta PW; Hecht SS; Sturla SJ; Balbo S Anal. Chem. 2015, 87, 11706–11713. [PubMed: 26509677]

(14). Liu S; Wang YS Anal. Chem. 2013, 85, 6732–6739. [PubMed: 23789926]

(15). Villalta PW; Balbo S Int. J. Mol. Sci. 2017, 18, 1870.

(16). Bizanek R; McGuinness BF; Nakanishi K; Tomasz M Biochemistry 1992, 31, 3084–3091. [PubMed: 1554696]

(17). Douki T; Laporte G; Cadet J Nucleic Acids Res. 2003, 31, 3134–3142. [PubMed: 12799441]

(18). Chen HJC; Chang YL; Teng YC; Hsiao CF; Lin TS Anal. Chem. 2017, 89, 13082–13088. [PubMed: 29172486]

(19). Paz MM; Ladwa S; Champeil E; Liu Y; Rockwell S; Boamah EK; Bargonetti J; Callahan J; Roach J; Tomasz M Chem. Res. Toxicol. 2008, 21, 2370–2378. [PubMed: 19053323]

(20). Stornetta A; Villalta PW; Gossner F; Wilson WR; Balbo S; Sturla SJ Chem. Res. Toxicol. 2017, 30, 830–839. [PubMed: 28140568]

(21). Le Pla RC; Ritchie KJ; Henderson CJ; Wolf CR; Harrington CF; Farmer PB Chem. Res. Toxicol. 2007, 20, 1177–1182. [PubMed: 17636892]

(22). Zdrowowicz M; Wityk P; Michalska B; Rak J Org. Biomol. Chem. 2016, 14, 9312–9321. [PubMed: 27714178]

(23). Cheng G; Shi Y; Sturla SJ; Jalas JR; McIntee EJ; Villalta PW; Wang M; Hecht SS Chem. Res. Toxicol. 2003, 16, 145–152. [PubMed: 12588185]

(24). Mohamed D; Mowaka S; Thomale J; Linscheid MW Chem. Res. Toxicol. 2009, 22, 1435–1446. [PubMed: 19621941]

(25). Hu CW; Chen CM; Ho HH; Chao MR Anal. Bioanal. Chem. 2012, 402, 1199–1208. [PubMed: 22094591]
(26). Chao MR; Yen CC; Hu CW Free Radical Biol. Med. 2008, 44, 464–473. [PubMed: 17983606]
(27). Chang YJ; Cooke MS; Hu CW; Chao MR Arch. Toxicol. 2018, 92, 2665–2680. [PubMed: 29943112]
(28). Wu JH; Jones NJ Methods Mol. Biol. 2012, 817, 165–181. [PubMed: 22147573]
(29). Kiakos K; Hartley JM; Hartley JA Methods Mol. Biol. 2010, 613, 283–302. [PubMed: 19997891]
(30). Ersson C; Moller P; Forchhammer L; Loft S; Azqueta A; Godschalk RW; van Schooten FJ; Jones GD; Higgins JA; Cooke MS; et al. Mutagenesis 2013, 28, 279–286. [PubMed: 23446176]
(31). Wang Y; Wang Y Anal. Chem. 2003, 75, 6306–6313. [PubMed: 14616015]
(32). Yun BH; Geacintov NE; Shafirovich V Chem. Res. Toxicol. 2011, 24, 1144–1152. [PubMed: 21513308]
(33). Baskerville-Abraham IM; Boysen G; Troutman JM; Mutlu E; Collins L; deKrafft KE; Lin W; King C; Chaney SG; Swenberg JA Chem. Res. Toxicol. 2009, 22, 905–912. [PubMed: 19323581]
(34). Chaw YF; Crane LE; Lange P; Shapiro R Biochemistry 1980, 19, 5525–5531. [PubMed: 7459328]
(35). Nejad MI; Johnson KM; Price NE; Gates KS Biochemistry 2016, 55, 7033–7041. [PubMed: 27992994]
(36). Grillari J; Katinger H; Voglauer R Nucleic Acids Res. 2007, 35, 7566–7576. [PubMed: 18083760]
(37). Chen H; Yao L; Brown C; Rizzo CJ; Turesky RJ Anal. Chem. 2019, 91, 7403–7410. [PubMed: 31055913]
(38). Ravanat JL; Douki T; Duez P; Gremaud E; Herbert K; Hofer T; Lasserre L; Saint-Pierre C; Favier A; Cadet J Carcinogenesis 2002, 23, 1911–1918. [PubMed: 12419840]
(39). Nakamura J; La DK; Swenberg JA J. Biol. Chem. 2000, 275, 5323–5328. [PubMed: 10681505]
(40). Gedik CM; Collins A FASEB J. 2005, 19, 82–84. [PubMed: 15533950]
(41). Singh RK; Kumar S; Prasad DN; Bhardwaj TR Eur. J. Med. Chem. 2018, 151, 401–433. [PubMed: 29649739]
(42). Gates KS Chem. Res. Toxicol. 2009, 22, 1747–1760. [PubMed: 19757819]
(43). Florea-Wang D; Pawlowicz AJ; Sinkkonen J; Kronberg L; Vilpo J; Hovinen J Chem. Biodiversity 2009, 6, 1002–1013.
(44). Ceccaldi R; Sarangi P; D'Andrea AD Nat. Rev. Mol. Cell Biol. 2016, 17, 337–349. [PubMed: 27145721]

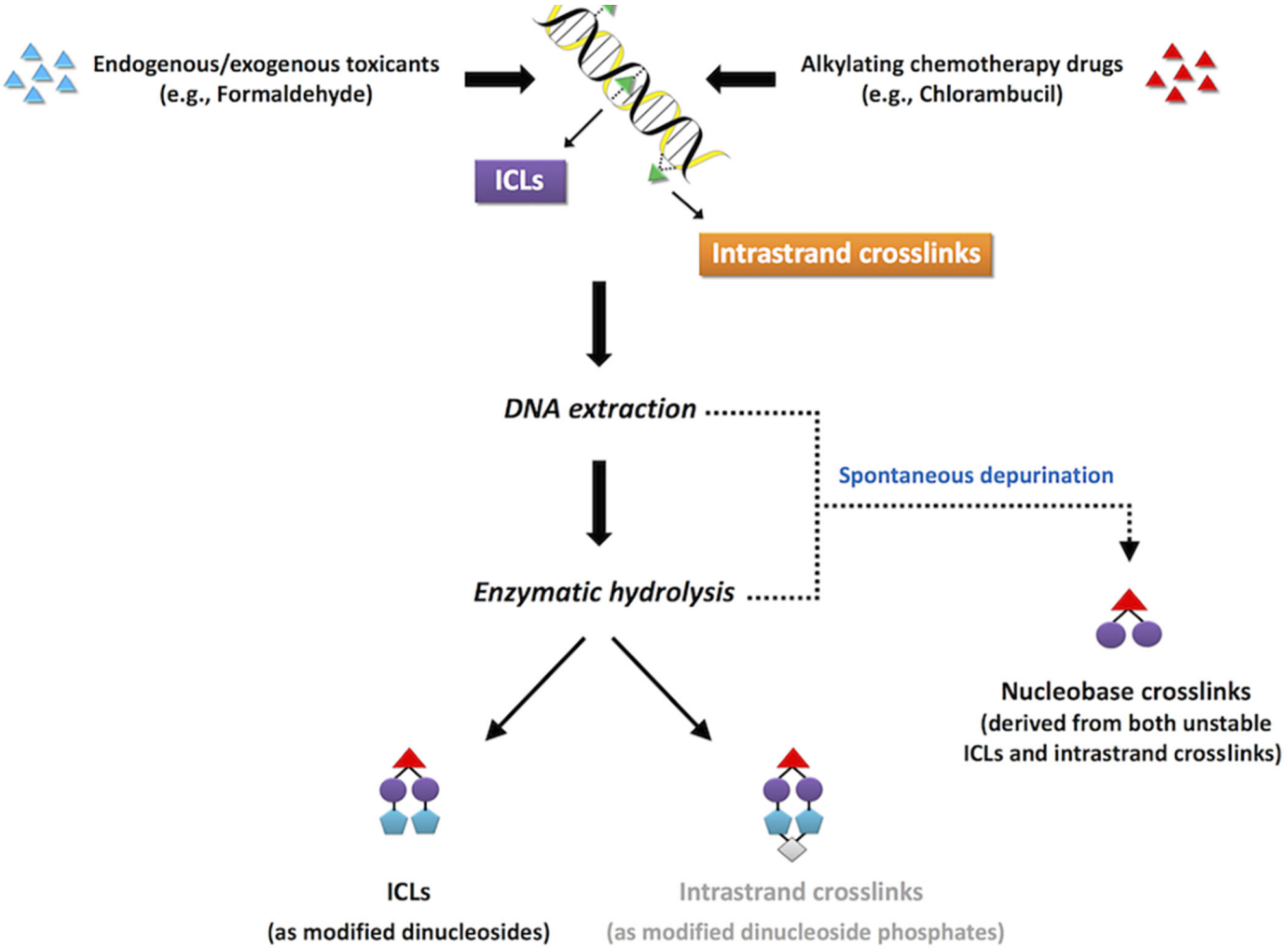


Figure 1.

Concepts underlying the DNA crosslinkomic approach. Both exogenous and endogenous sources (e.g., formaldehyde and alkylating chemotherapy drug) could induce DNA ICLs and intrastrand crosslinks. In the proposed DNA crosslinkomic approach, a DNA sample was isolated and enzymatically hydrolyzed to selectively release ICLs (as modified dinucleosides) and intrastrand crosslinks (as modified dinucleoside phosphates). During sample preparation, some purine nucleosides, modified at certain nucleobase positions, are unstable and may undergo spontaneous depurination, resulting in the formation of nucleobase crosslinks. This study aimed to comprehensively measure both ICLs and nucleobase crosslinks.

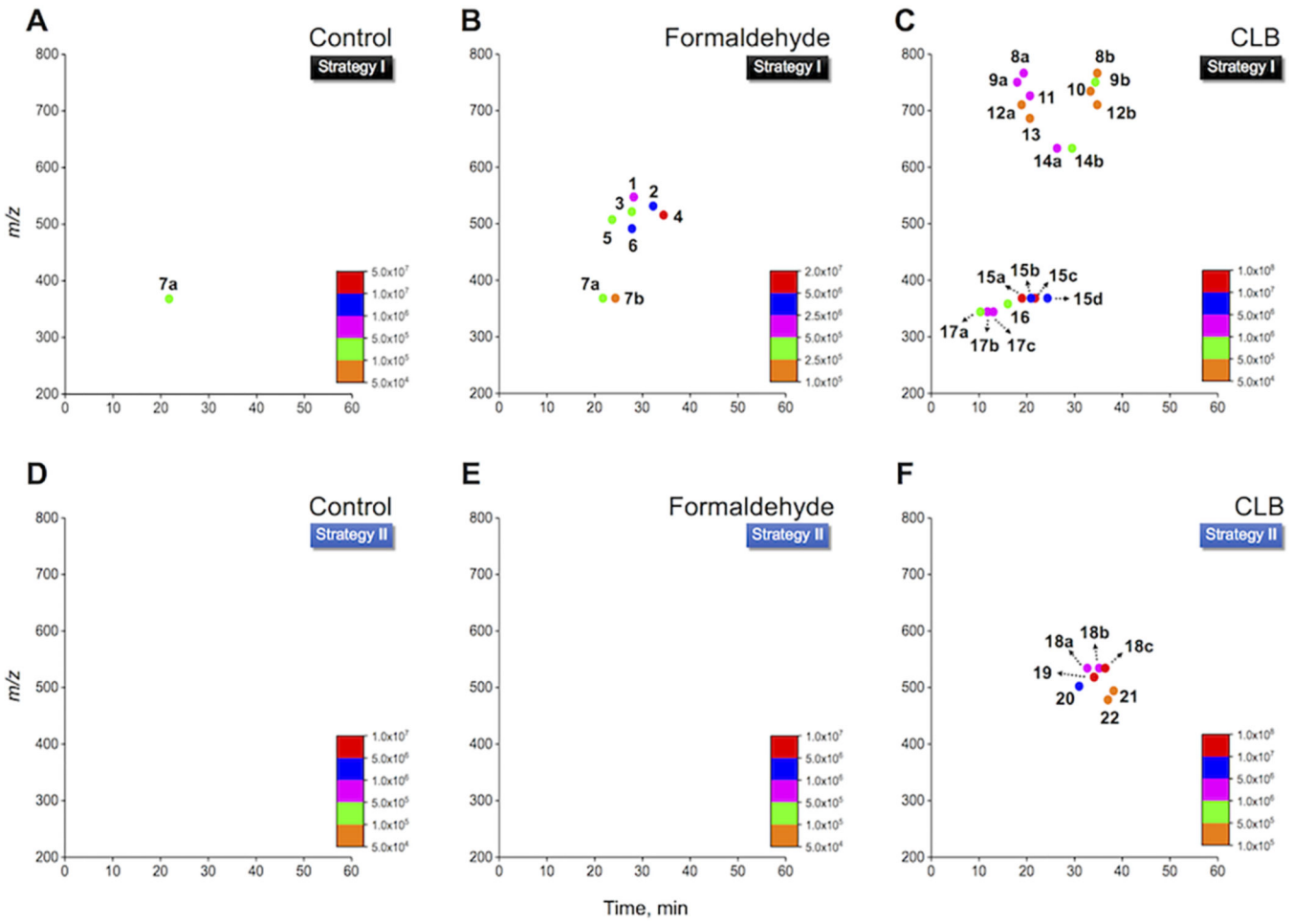
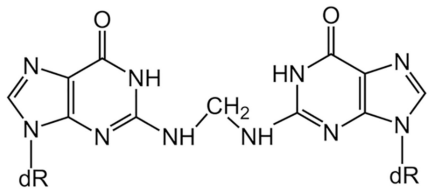
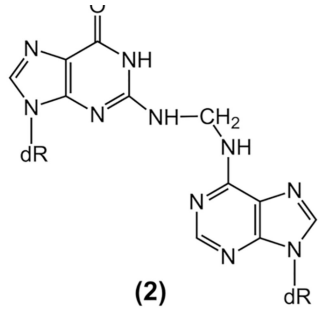


Figure 2.

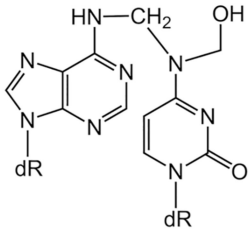
DNA crosslinkome maps obtained from (A) control CT-DNA, (B) formaldehyde-treated CT-DNA, and (C) CLB-treated CT-DNA measured using Strategy I, whereas Strategy II was used to analyze (D) control CT-DNA, (E) formaldehyde-treated CT-DNA, and (F) CLB-treated CT-DNA. Detected ions are summarized in Table S1 for Ions 1–7, Table S2 for Ions 8–17, and Table S3 for Ions 18–22.



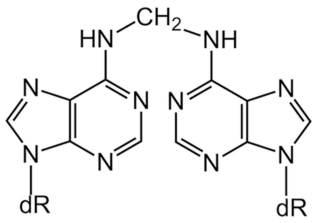
(1)
dG-CH_2-dG
Theoretical mass: 546.1935
Formula: $C_{21}H_{26}N_{10}O_8$



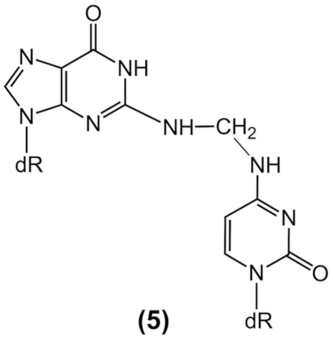
(2)
dG-CH_2-dA
Theoretical mass: 530.1986
Formula: $C_{21}H_{26}N_{10}O_7$



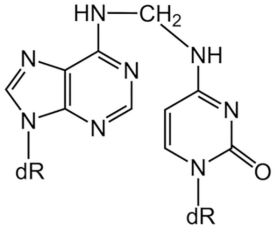
(3)
dA-CH_2-hmdC
Theoretical mass: 520.2030
Formula: $C_{21}H_{28}N_8O_8$



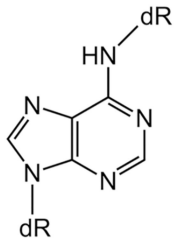
(4)
dA-CH_2-dA
Theoretical mass: 514.2037
Formula: $C_{21}H_{26}N_{10}O_6$



(5)
dG-CH_2-dC
Theoretical mass: 506.1874
Formula: $C_{21}H_{26}N_{10}O_6$



(6)
dA-CH_2-dC
Theoretical mass: 490.1924
Formula: $C_{20}H_{26}N_8O_7$



(7)
dA-AP
Theoretical mass: 367.1492
Formula: $C_{15}H_{21}N_5O_6$

Figure 3.
Proposed chemical structures of formaldehyde-induced DNA ICLs (Ions 1–7) in CT-DNA, measured as modified dinucleosides using Strategy I.

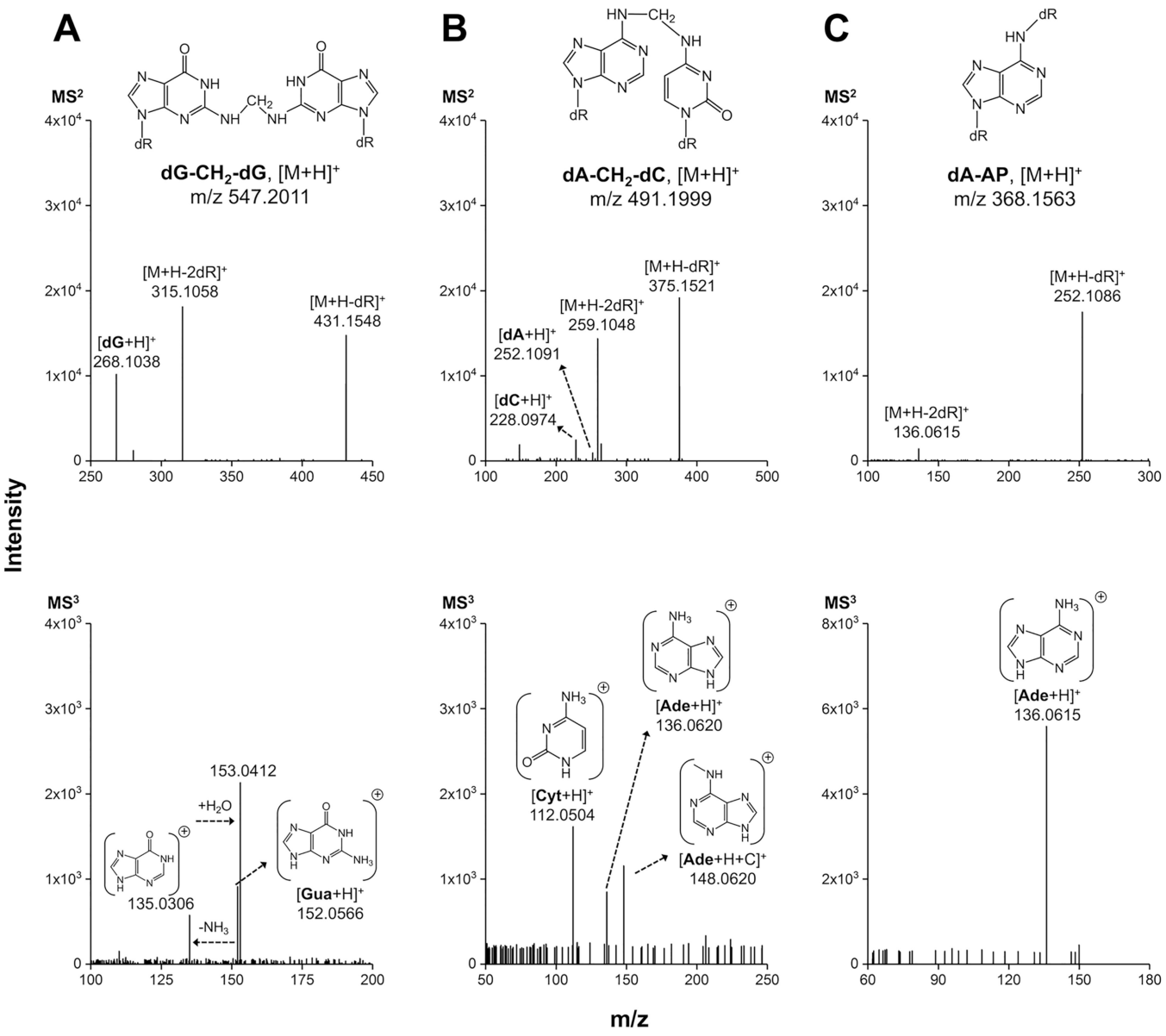


Figure 4.

Product ion spectra of formaldehyde-induced DNA crosslinks measured as modified dinucleosides using Strategy I: (A) dG-CH_2-dG as Ion 1, (B) dA-CH_2-dC as Ion 6, and (C) dA-AP as Ion 7.

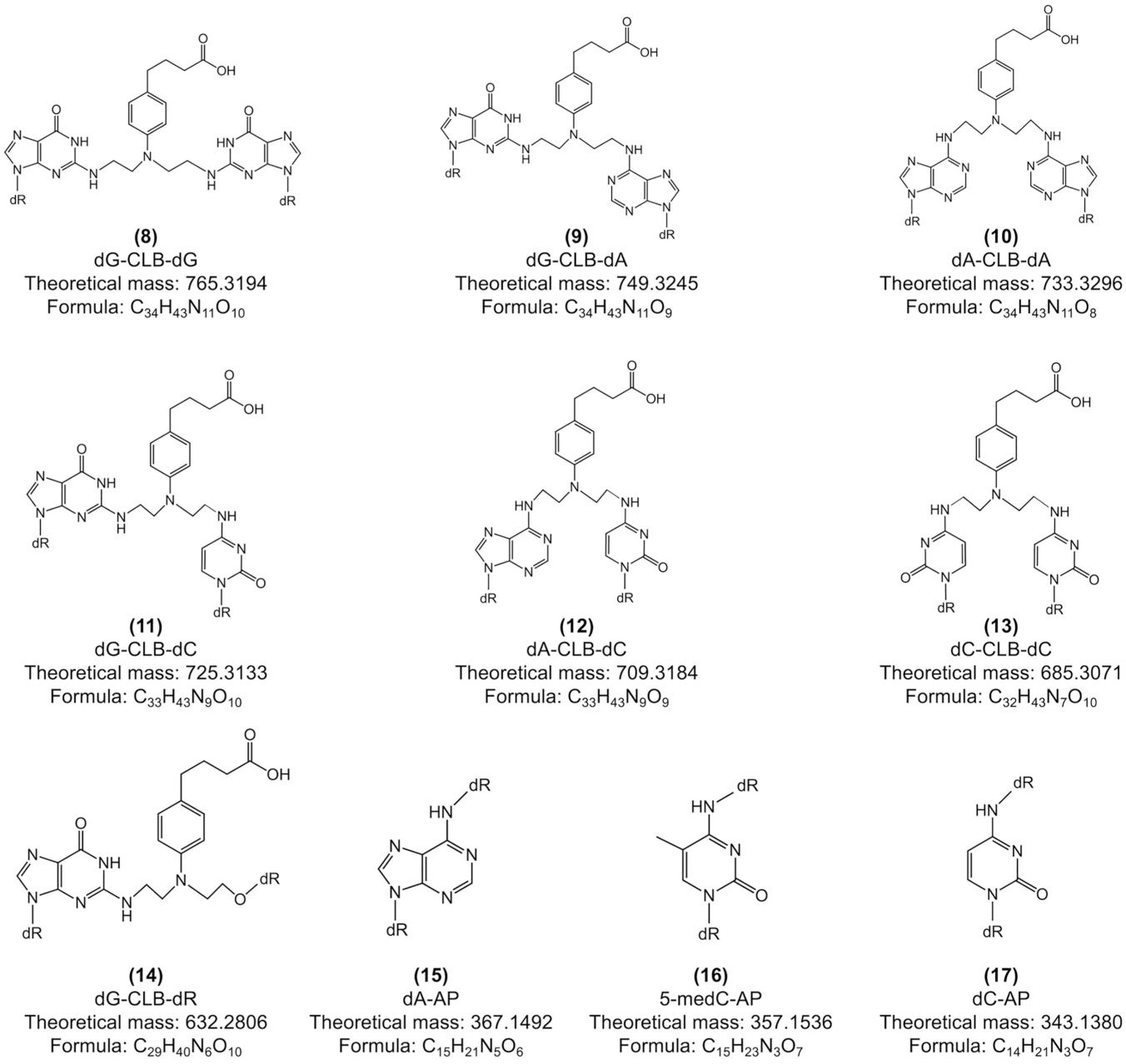


Figure 5.

Proposed chemical structures of CLB-induced DNA ICLs (Ions 8–17) in CT-DNA, measured as modified dinucleosides using Strategy I.

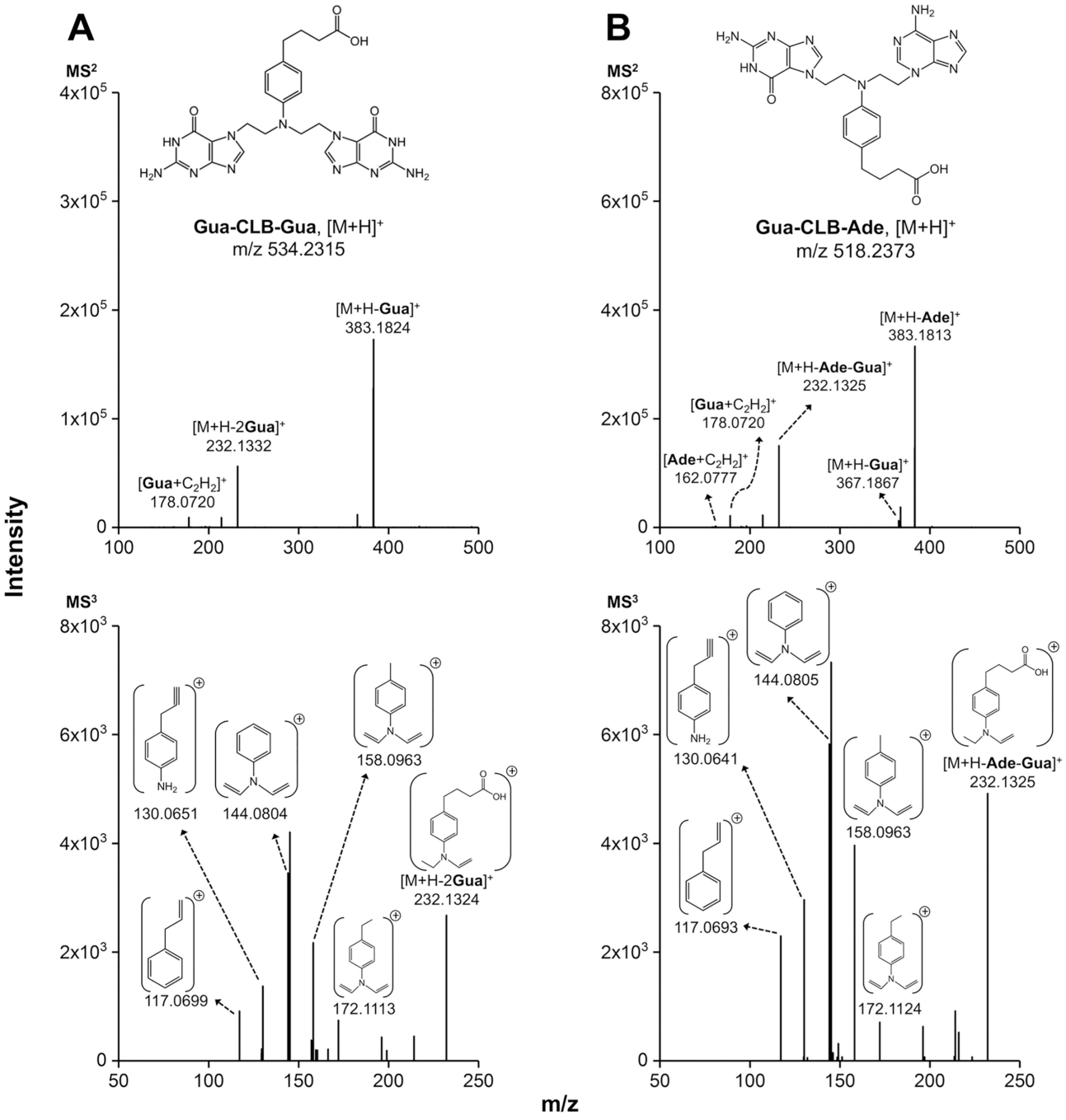


Figure 6.

Product ion spectra of CLB-induced DNA crosslinks measured as modified dinucleobases using Strategy II: (A) Gua-CLB-Gua as Ion 18 and (B) Gua-CLB-Ade as Ion 19.