HHS Public Access

Author manuscript

Biochemistry. Author manuscript; available in PMC 2020 February 04.

Published in final edited form as:

Biochemistry. 2018 November 13; 57(45): 6391–6394. doi:10.1021/acs.biochem.8b01023.

Characterization of Natural Colibactin–Nucleobase Adducts by Tandem Mass Spectrometry and Isotopic Labeling. Support for DNA Alkylation by Cyclopropane Ring Opening

Mengzhao Xue[†], **Emilee Shine**[‡,§], **Weiwei Wang**[‖,⊥], **Jason M. Crawford**[*,†,‡,§], **Seth B. Herzon**[*,†,#]

[†]Department of Chemistry, Yale University, New Haven, Connecticut 06520, United States

[‡]Chemical Biology Institute, Yale University, West Haven, Connecticut 06516, United States

[§]Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, Connecticut 06536, United States

[‖]Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, P.O. Box 208114, New Haven, Connecticut 06520, United States

[⊥]W. M. Keck Biotechnology Resource Laboratory, Yale University School of Medicine, 300 George Street, New Haven, Connecticut 06510, United States

[#]Department of Pharmacology, Yale University School of Medicine, New Haven, Connecticut 06520, United States

Abstract

Colibactins are genotoxic secondary metabolites whose biosynthesis is encoded in the *clb* gene cluster harbored by certain strains of gut commensal *Escherichia coli.* Using synthetic colibactin analogues, we previously provided evidence that colibactins alkylate DNA by addition of a nucleotide to an electrophilic cyclopropane intermediate. However, natural colibactin–nucleobase adducts have not been identified, to the best of our knowledge. Here we present the first identification of such adducts, derived from treatment of pUC19 DNA with *clb*$^+$ *E. coli.* Previous biosynthetic studies established cysteine and methionine as building blocks in colibactin biosynthesis; accordingly, we used cysteine (Δ*cys*E) and methionine (Δ*met*A) auxotrophic strains cultured in media supplemented with L-[U-^{13}C]Cys or L-[U-^{13}C]Met to facilitate the identification of nucleobases bound to colibactins. Using MS^2 and MS^3 analysis, in conjunction with the known oxidative instability of colibactin cyclopropane-opened products, we were able to characterize adenine adducts derived from cyclopropane ring opening. This study provides the first reported detection of nucleobase adducts derived from *clb*$^+$ *E. coli* and lends support to our earlier model suggesting DNA alkylation by addition of a nucleotide to an electrophilic cyclopropane.

[*]**Corresponding Authors**: seth.herzon@yale.edu, jason.crawford@yale.edu.

Author Contributions

S.B.H. wrote the manuscript.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.8b01023.

Figures S1–S13, Tables S1–S3, detailed experimental procedures, and characterization data for all new compounds (PDF)

The authors declare no competing financial interest.

Colibactins are secondary metabolites encoded in a hybrid nonribosomal peptide synthetase–polyketide synthase (NRPS–PKS) gene cluster, termed *clb*, that is harbored by certain strains of gut commensal *Escherichia coli*.[1–5]

The presence of the *clb* cluster is epidemiologically correlated to colorectal cancer formation in humans, and studies suggest colibactins are genotoxic and cause tumors in mouse models.

We previously provided evidence that colibactin genotoxicity derives in part from addition of a nucleotide to an electrophilic cyclopropane,[6] a mechanism of DNA damage established for several classes of natural products.[7,8] Because natural colibactins have eluded isolation, this conclusion was based on studies of synthetic colibactin analogues such as **1**–**3** (Scheme 1A). First, linearized pBR322 plasmid DNA was extensively degraded following treatment with nanomolar concentrations of synthetic colibactin analogue **1**, which bears the putative electrophilic cyclopropane residue. Second, and consistent with the cyclopropane behaving as a DNA electrophile, dimer **2** was found to form DNA interstrand cross-links. Finally, construct **3**, which is identical to **1** save for conversion of the cyclopropane to a *geminal*-dimethyl substituent, did not lead to detectable levels of DNA damage (≤500 μM concentration), as expected if the cyclopropane were a DNA-reactive functionality.

Perhaps the strongest evidence implicating the cyclopropane as a reactive locus underlying the genotoxicity of *clb*$^+$ *E. coli* was obtained through studies of the resistance enzyme ClbS. ClbS is encoded in the *clb* cluster and was shown to be essential for bacterial viability.[9] We demonstrated that purified ClbS cleaved the cyclopropane residue in synthetic colibactin **4**, ultimately resulting in formation of 3-hydroxytetrahydrofuran **7** (Scheme 1B).[10] The formation of **7** was shown to proceed by ClbS-catalyzed cyclopropane hydrolysis (**4** → **5**), aerobic oxidation (**5** → **6**), and cyclization with concomitant reduction of the alkyl hydrogen peroxide (**6** → **7**). This study established that the bacteria evolved a mechanism to eliminate self-toxicity deriving from the reactivity of the cyclopropane.

Despite these advances, the identification of natural colibactin–nucleobase adducts has not been described, to the best of our knowledge. In a recent study, *clb*$^+$ *E. coli* were demonstrated to cross-link exogenous DNA.[11] This suggested to us the possibility of characterizing natural colibactin–nucleobase adducts directly from bacterial cultures. To achieve this, here we conducted tandem mass spectrometry (MS) analysis of the products formed after incubation of linearized pUC19 plasmid DNA with *clb*$^+$ *E. coli* BW25113. We conducted parallel assays using a cysteine auxotroph (Δ*cys*E) and a methionine auxotroph (Δ*met*A)[12] cultured in media supplemented with L-[U-^{13}C]Cys or L-[U-^{13}C]Met, respectively. Biochemical studies have established that the thiazole and aminocyclopropane residues of colibactins are derived from cysteine[13] and methionine via SAM,[13–15] respectively. Thus, products derived from *clb* metabolites were expected to be mass-shifted by 3 units for each cysteine or 4 units for each methionine in these auxotrophs, facilitating their identification. *clb*$^-$ *E. coli* were used as a negative control.

Linearized pUC19 DNA and bacteria were incubated in M9 medium for 4.5 h at 37 °C. The bacteria were separated by centrifugation, and the DNA was isolated and analyzed by denaturing gel electrophoresis. As shown in Figure 1, DNA was cross-linked upon being

exposed to *clb*$^+$, Δ*cys*E, or Δ*met*A *E. coli* but not upon being exposed to *clb*$^-$ *E. coli*, as expected. The cross-linked DNA was digested with the Nucleoside Digestion Mix (New England Biolabs) and analyzed by liquid chromatography and tandem MS.

Prominent peaks at *m/z* 522.1668, 538.1618, 540.1775, and 556.1722 (*z* = 1) and *m/z* 261.5871, 269.5844, 270.5925, and 278.5899 (*z* = 2) were identified in DNA treated with *clb*$^+$ *E. coli*. These peaks were mass-shifted by +3 or +4 units (*z* = 1) or +1.5 or +2 units (*z* = 2) in the Δ*cys*E/L-[U-^{13}C]Cys or Δ*met*A/L-[U-^{13}C]Met cultures, respectively. The +3 or +1.5 mass shift in the products derived from the Δ*cys*E culture indicates the presence of only one thiazole ring. These observed masses fit the proposed ion structures **8**–**11** within 1.5 ppm of error (Figure 2 and Table 1).

The connectivity of **8**–**11** (in particular, the location of the adenine base) was established by extensive MS^2 and MS^3 analysis in conjunction with the mass shifts in the auxotrophic strains anticipated on the basis of the known label origins of the thiazole and cyclopropane residues.[13,14]

Thus, daughter ions **15** and **19** were observed in the MS spectra of **10** (Scheme 2). Ion **14** was observed as a daughter ion in the MS and MS^2 spectra of **10**, while ions **12** and **13** were observed only in the MS^2 spectrum of **10**. Thiazole **14** was mass-shifted by +3 units in the Δ*cys*E culture but did not change in the Δ*met*A culture, consistent with the known biosynthesis of the thiazole residues from cysteine. Ions **15** and **19** were mass-shifted by +4 units in the Δ*met*A culture but did not shift in the Δ*cys*E culture, consistent with the derivation of the cyclopropane from labeled methionine. The masses of adenine·H^+ adducts were detected in the MS^2 spectrum of unlabeled **10** and its Cys and Met-labeled isotopologs (error of <3 ppm). Indeed, the exact masses of **15** and **19** support incorporation of adenine, and their attendant +4 mass shift in the Δ*met*A auxotroph allows us to associate the adenine residue with the region that contained the cyclopropane. Ion **12** further fragmented to **16** and **17**, and consistent with their structures, all five of the ions (**10**, **12**, **13**, **16**, and **17**) were mass-shifted by +3 and +4 units in the Cys and Met auxotrophs, respectively. The fragmentation of **17** to **16** results in a loss of 43 mass units. This difference is consistent with loss of a fragment containing a nitrogen atom and supports the location of the hydroxyl group in **17** and in its parent ions. Future studies will focus on determining the site of formation of the bond to adenine (e.g., N7, C6–NH_2).

Previously, it has been shown that advanced colibactins contain a two-carbon spacer between the thiazole rings.[16] This two-carbon spacer is derived from an α-aminomalonate residue and has been shown to be essential for genotoxic effects.[17,18] The identification of colibactin–nucleobase monoadducts **8**–**11** derived from the cross-linked precursor product suggests a potential role for this two-carbon spacer in cross-linking and genotoxicity, as the spacer and one of the thiazole rings are lacking in the detectable nuclease digestion products. Additionally, the cellular role of the electrophilicity of the lactam of metabolites resembling cyclopropane ring-opened **6** remains unknown. Further studies will be required to fully elucidate these points.

In summary, we have described the first structural evidence of the production of natural colibactin–nucleobase adducts. This study lends further support to our earlier work suggesting the cyclopropane as a DNA-reactive locus. Further studies will focus on elucidating the second site of reactivity leading to formation of DNA cross-links.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

Financial support from the National Institutes of Health (1R01CA215553; SIG Grants 1S10OD019967 and 1S10ODOD018034) is gratefully acknowledged.

REFERENCES

(1). Trautman EP, and Crawford JM (2016) Linking biosynthetic gene clusters to their metabolites via pathway-targeted molecular networking. Curr. Top. Med. Chem 16, 1705–1716. [PubMed: 26456470]

(2). Balskus EP (2015) Colibactin: Understanding an elusive gut bacterial genotoxin. Nat. Prod. Rep 32, 1534–1540. [PubMed: 26390983]

(3). Taieb F, Petit C, Nougayrede JP, and Oswald E (2016) The enterobacterial genotoxins: Cytolethal distending toxin and colibactin. EcoSal Plus 7, n/a DOI: 10.1128/ecosalplus.ESP-0008-2016.

(4). Healy AR, and Herzon SB (2017) Molecular basis of gut microbiome-associated colorectal cancer: A synthetic perspective. J. Am. Chem. Soc 139, 14817–14824. [PubMed: 28949546]

(5). Faïs T, Delmas J, Barnich N, Bonnet R, and Dalmasso G (2018) Colibactin: More than a new bacterial toxin. Toxins 10, 151.

(6). Healy AR, Nikolayevskiy H, Patel JR, Crawford JM, and Herzon SB (2016) A mechanistic model for colibactin-induced genotoxicity. J. Am. Chem. Soc 138, 15563–15570. [PubMed: 27934011]

(7). Boger DL, and Garbaccio RM (1999) Shape-dependent catalysis: Insights into the source of catalysis for the CC-1065 and duocarmycin DNA alkylation reaction. Acc. Chem. Res 32, 1043–1052.

(8). Tichenor MS, and Boger DL (2008) Yatakemycin: Total synthesis, DNA alkylation, and biological properties. Nat. Prod. Rep 25, 220–226. [PubMed: 18389136]

(9). Bossuet-Greif N, Dubois D, Petit C, Tronnet S, Martin P, Bonnet R, Oswald E, and Nougayrede JP (2016) Escherichia coli ClbS is a colibactin resistance protein. Mol. Microbiol 99, 897–908. [PubMed: 26560421]

(10). Tripathi P, Shine EE, Healy AR, Kim CS, Herzon SB, Bruner SD, and Crawford JM (2017) ClbS is a cyclopropane hydrolase that confers colibactin resistance. J. Am. Chem. Soc 139, 17719–17722. [PubMed: 29112397]

(11). Bossuet-Greif N, Vignard J, Taieb F, Mirey G, Dubois D, Petit C, Oswald E, and Nougayrede JP (2018) The colibactin genotoxin generates DNA interstrand cross-links in infected cells. mBio 9, e02393–17. [PubMed: 29559578]

(12). Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, and Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. Mol. Syst. Biol 2, 2006.0008.

(13). Vizcaino MI, and Crawford JM (2015) The colibactin warhead crosslinks DNA. Nat. Chem 7, 411–417. [PubMed: 25901819]

(14). Bian X, Plaza A, Zhang Y, and Müller R (2015) Two more pieces of the colibactin genotoxin puzzle from *Escherichia coli* show incorporation of an unusual 1-aminocyclopropanecarboxylic acid moiety. Chem. Sci 6, 3154–3160. [PubMed: 28706687]

(15). Zha L, Jiang Y, Henke MT, Wilson MR, Wang JX, Kelleher NL, and Balskus EP (2017) Colibactin assembly line enzymes use *S*-adenosylmethionine to build a cyclopropane ring. Nat. Chem. Biol 13, 1063. [PubMed: 28805802]

(16). Li ZR, Li J, Gu JP, Lai JY, Duggan BM, Zhang WP, Li ZL, Li YX, Tong RB, Xu Y, Lin DH, Moore BS, and Qian PY (2016) Divergent biosynthesis yields a cytotoxic aminomalonate-containing precolibactin. Nat. Chem. Biol 12, 773–775. [PubMed: 27547923]

(17). Nougayrède J-P, Homburg S, Taieb F, Boury M, Brzuszkiewicz E, Gottschalk G, Buchrieser C, Hacker J, Dobrindt U, and Oswald E (2006) *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. Science 313, 848–851. [PubMed: 16902142]

(18). Brachmann AO, Garcie C, Wu V, Martin P, Ueoka R, Oswald E, and Piel J (2015) Colibactin biosynthesis and biological activity depend on the rare aminomalonyl polyketide precursor. Chem. Commun 51, 13138–13141.

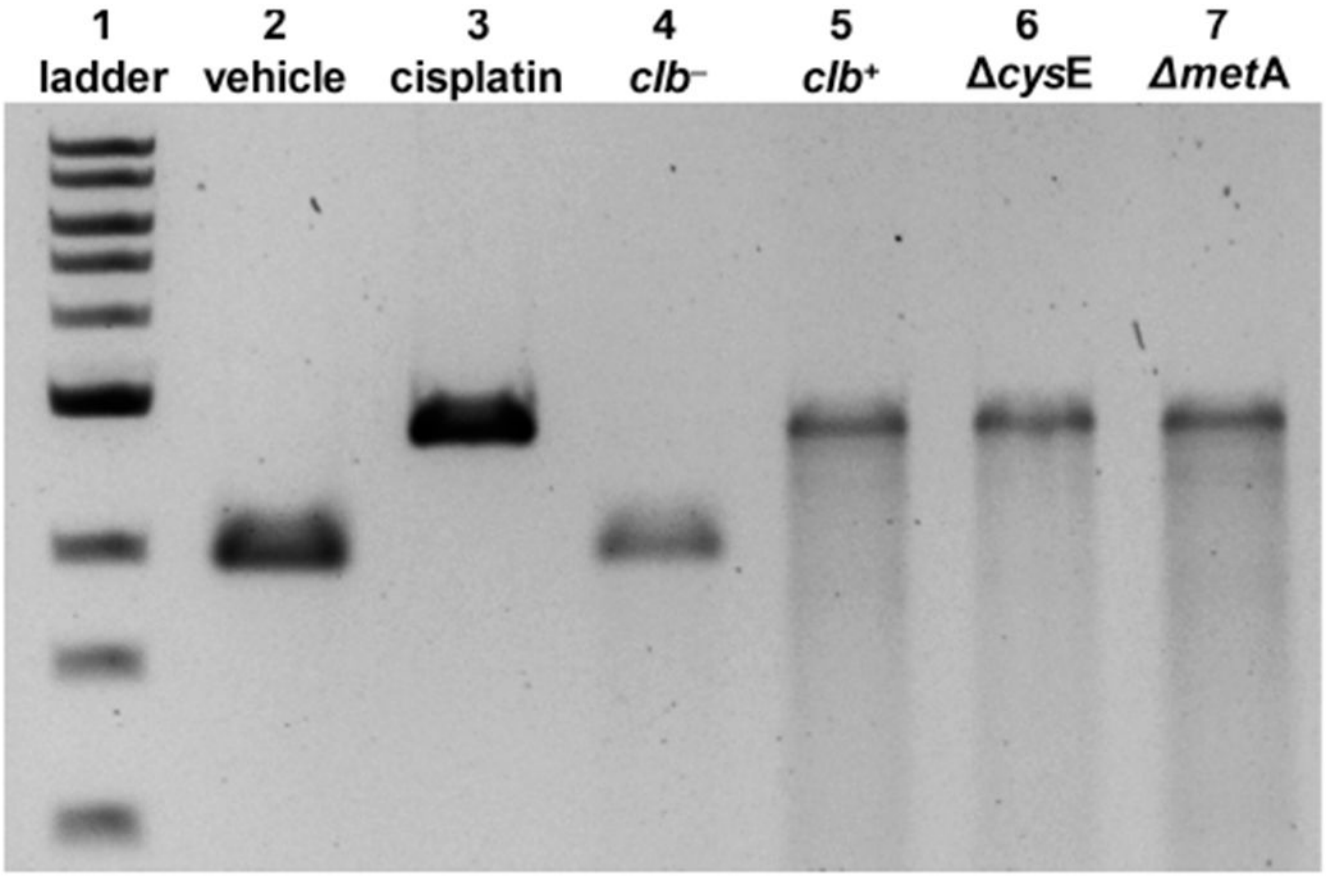


Figure 1.

DNA cross-linking assays employing linearized pUC19 DNA and *E. coli* variants. Cisplatin was used as a positive control: DNA ladder (lane 1), no treatment (lane 2), cisplatin (100 μM, lane 3), *clb*$^{-}$ BW25113 *E. coli* (lane 4), *clb*$^{+}$ BW25113 *E. coli* (lane 5), Δ*cys*E *clb*$^{+}$ BW25113 *E. coli* (lane 6), and Δ*met*A *clb*$^{+}$ BW25113 *E. coli* (lane 7). Conditions for lanes 2 and 3: linearized pUC19 DNA (31 μM in base pairs), pH 5 sodium citrate buffer (10 mM), 4.5 h, 37 °C. Conditions for lanes 4 and 5: linearized pUC19 DNA, M9 medium, 4.5 h, 37 °C. Conditions for lanes 6 and 7: linearized pUC19 DNA, modified M9 medium (containing L-[U-^{13}C]Cys or L-[U-^{13}C]Met for Cys and Met auxotrophs, respectively), 4.5 h, 37 °C. DNA was isolated and analyzed by denaturing agarose gel electrophoresis (90 V, 1.5 h).

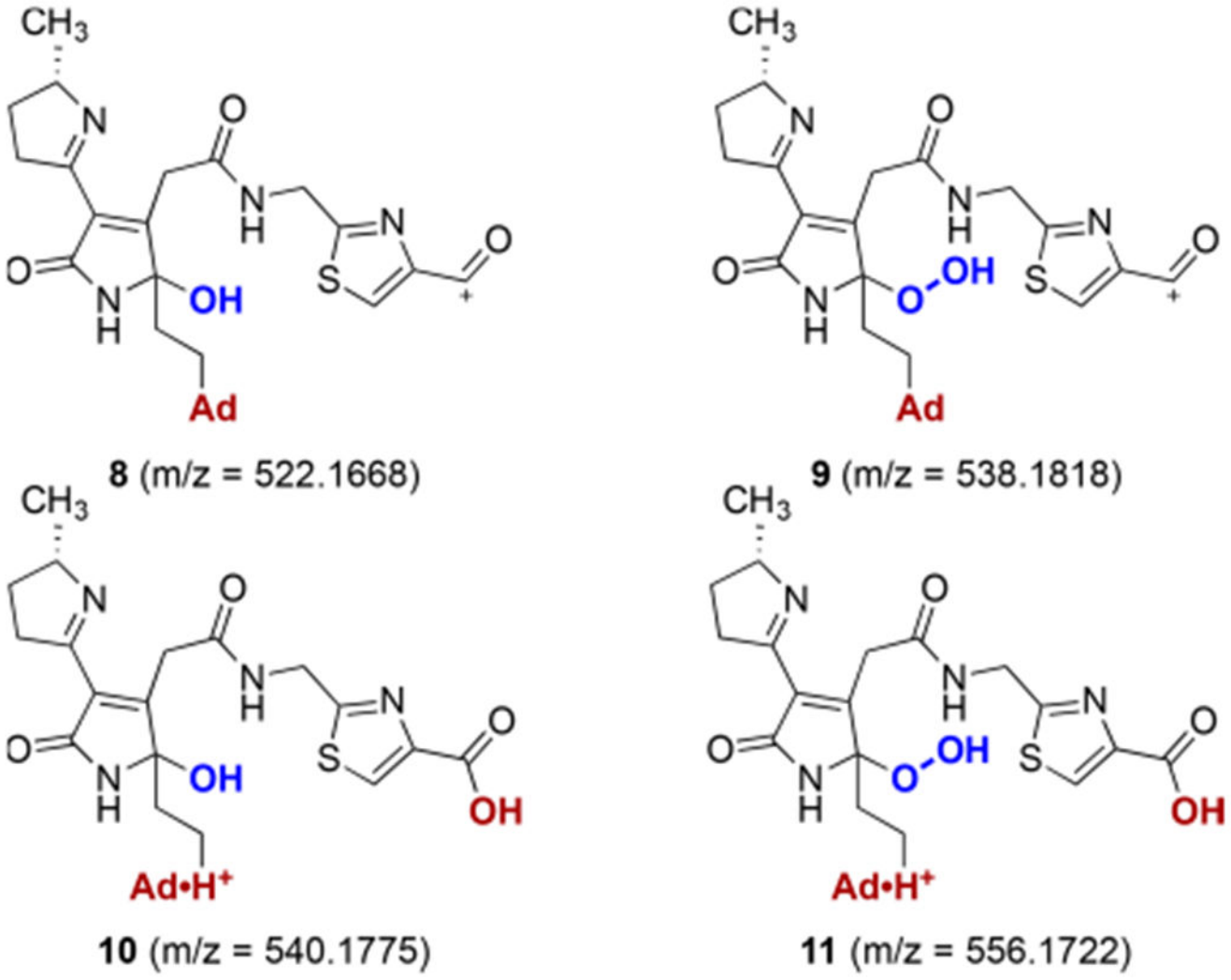
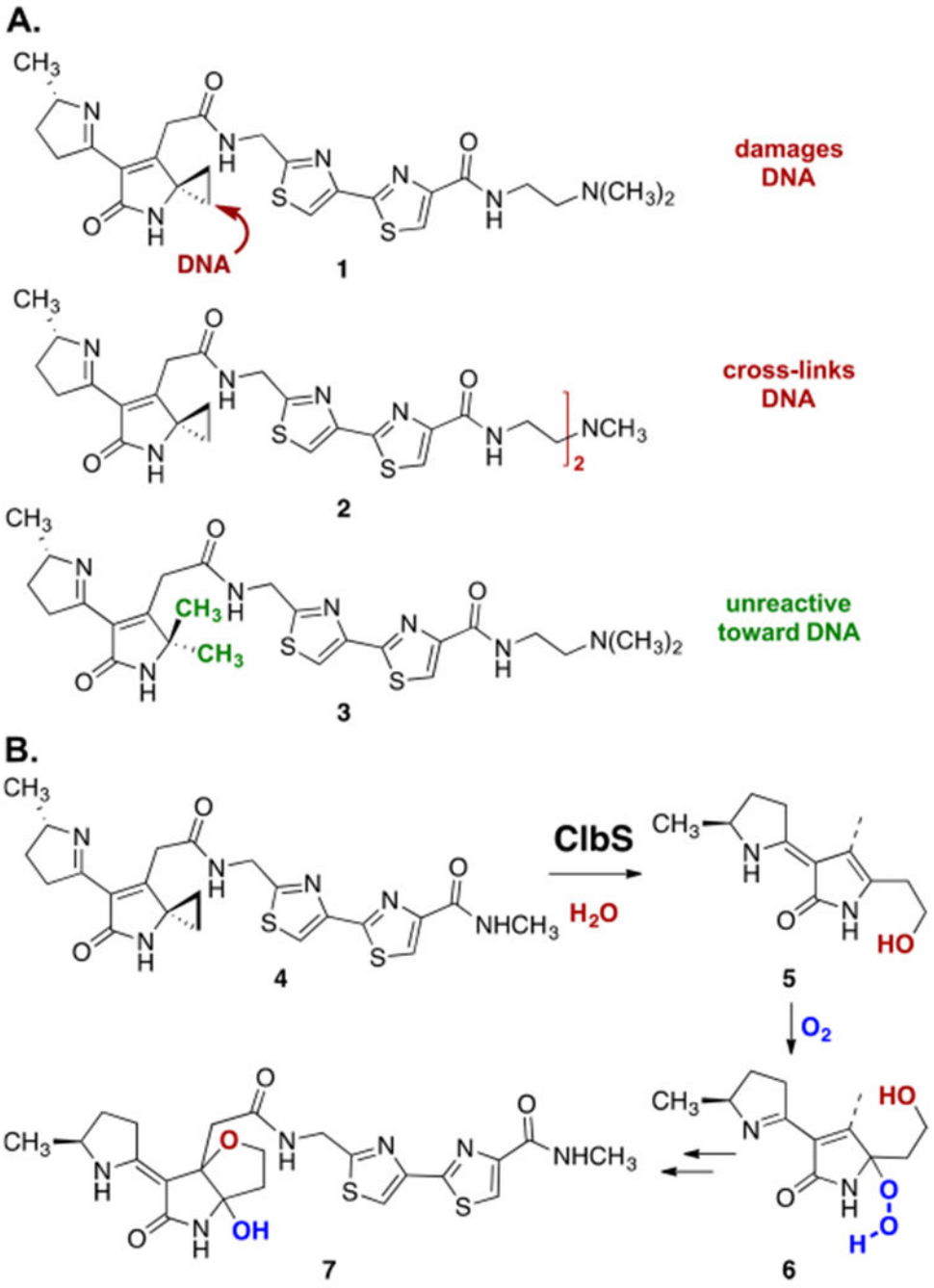


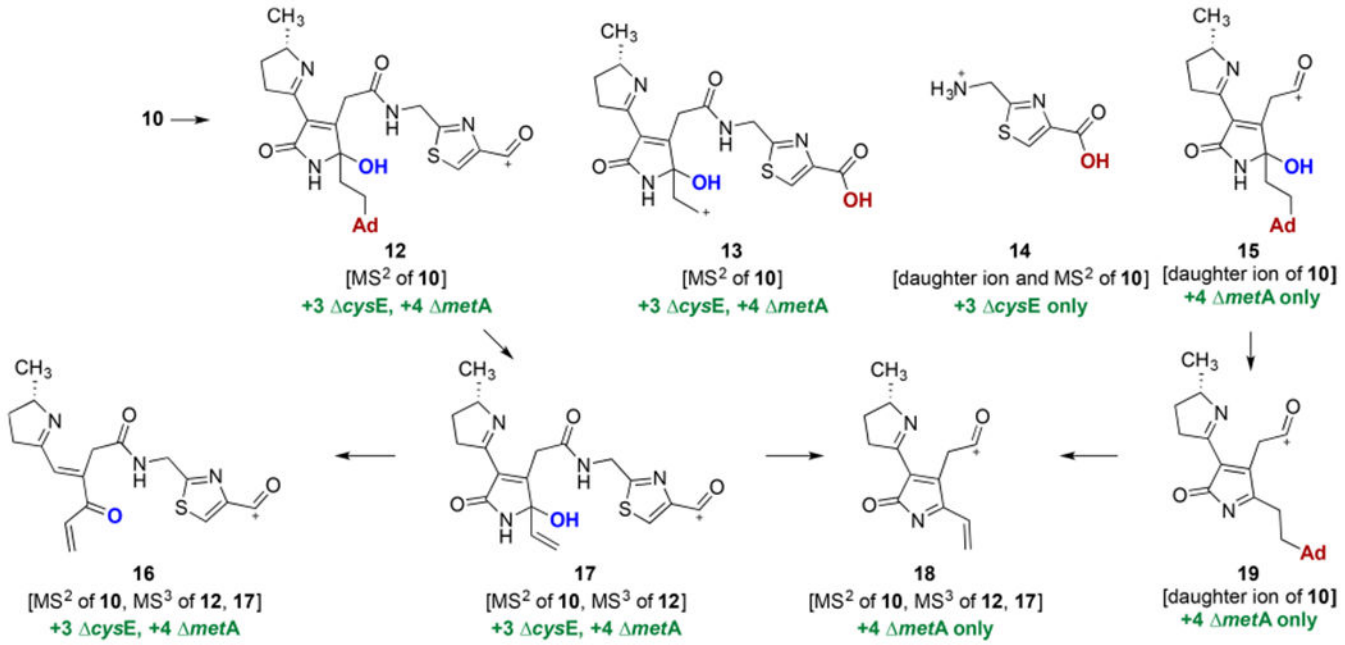
Figure 2.

Proposed structures of colibactin–adenine adducts derived from incubation of pUC19 DNA with *clb*$^{+}$ *E. coli.* Ad = adenine.



Scheme 1.

(A) Structures of Synthetic Colibactins 1–3 and Their Reactivity toward DNA and (B) Pathway for ClbS-Mediated Conversion of 4 to Hydroxyfuran 7



Scheme 2.

Selected Fragmentation and Proposed Tandem MS Products Derived from 10

Table 1.

High-Resolution Mass Spectrometry Data of Colibactin–Nucleobase Adducts

strain	ion	z	exp	theo	error (ppm)
clb$^+$	**8**	1	522.1668	522.1666	0.38
	9	1	538.1618	538.1616	0.37
	10	1	540.1775	540.1772	0.56
	11	1	556.1722	556.1721	0.18
	8	2	261.5871	261.5870	0.38
	9	2	269.5844	269.5844	0.00
	10	2	270.5925	270.5922	1.11
	11	2	278.5899	278.5897	0.72
clb$^+$Δ*cys*E	[$^{13}C_3$]-**8**	1	525.1765	525.1767	0.29
	[$^{13}C_3$]-**9**	1	541.1714	541.1717	0.46
	[$^{13}C_3$]-**10**	1	543.1874	543.1873	0.28
	[$^{13}C_3$]-**11**	1	559.1822	559.1822	0.09
	[$^{13}C_3$]-**8**	2	263.0920	263.0920	0.10
	[$^{13}C_3$]-**9**	2	271.0897	271.0894	1.01
	[$^{13}C_3$]-**10**	2	272.0976	272.0972	1.38
	[$^{13}C_3$]-**11**	2	280.0947	280.0947	0.09
clb$^+$Δ*met*A	[$^{13}C_4$]-**8**	1	526.1804	526.1800	0.76
	[$^{13}C_4$]-**9**	1	542.1745	542.1750	0.92
	[$^{13}C_4$]-**10**	1	544.1907	544.1906	0.18
	[$^{13}C_4$]-**11**	1	560.1860	560.1855	0.89
	[$^{13}C_4$]-**8**	2	263.5938	263.5937	0.38
	[$^{13}C_4$]-**9**	2	271.5914	271.5911	1.10
	[$^{13}C_4$]-**10**	2	272.5992	272.5989	1.10
	[$^{13}C_4$]-**11**	2	280.5965	280.5964	0.36