MICROBIOLOGY

Revealing a microbial carcinogen

An *E. coli*–derived colibactin-DNA adduct is detected in intestinal tissues

By **Rachel M. Bleich**[1] *and* **Janelle C. Arthur**[1,2,3]

The microbiota in the human gastrointestinal system is predicted to produce hundreds of unique small molecules and secondary metabolites that may influence host health and disease (*1*). Many such molecules are produced by sophisticated multienzymatic assembly lines that are encoded by bacterial biosynthetic gene clusters. One class of molecules, colibactins, are produced from the gene cluster called the polyketide synthase (*pks*) island. The *pks* island occurs in certain strains of *Escherichia coli* and is prevalent in the microbiota of colorectal cancer (CRC) patients (*2*–*5*). However, despite more than a decade of research into the potential carcinogenic role of colibactin, little is known about its structure or mechanism of action. On page 709 of this issue, Wilson *et al.* (*6*) show that colibactin alkylates DNA in cultured cells and in vivo, forming covalent modifications known as DNA adducts. These colibactin-DNA adducts are chemical evidence of DNA damage and represent a detectable signature of exposure to colibactin. Misrepaired DNA adducts may generate mutations that contribute to colorectal tumorigenesis.

Colibactin was first described as an unknown product of a 54-kilobase genomic island that encodes a hybrid nonribosomal peptide synthetase–polyketide synthase gene cluster, the *pks* island, in some commensal and extraintestinal pathogenic *E. coli* strains (*2*). Exposure to *pks*$^+$ *E. coli* induces DNA double-strand breaks and an increased gene mutation frequency in mammalian cells in culture (*3*). This raised speculation that products of *pks* were microbial-derived genotoxins that could promote cancer. The tumorigenic potential of *pks* products was demonstrated in a study showing that *pks*$^+$ *E. coli* was abundant in colon tissue from CRC patients and promoted CRC in mouse models (*4*). This tumorigenic effect was later demonstrated in several other mouse models of CRC (*5*, *7*, *8*).

[1]*Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, NC, USA.* [2]*Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA.* [3]*Center for Gastrointestinal Biology and Disease, University of North Carolina, Chapel Hill, NC, USA. Email: janelle_arthur@med.unc.edu*

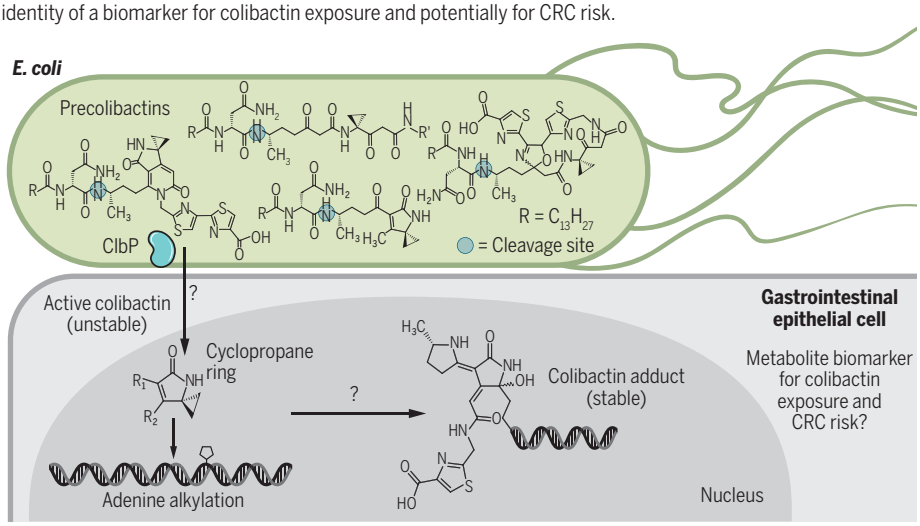
Because of its instability, the structure of colibactin has been elusive (*2*). Most previous work to determine the structure of colibactin has focused on identifying stable precursors using mutant strains of *E. coli* missing the colibactin-producing peptidase ClbP, which activates colibactin precursors by removing an *N*-myristoyl-D-asparagine "prodrug group" (*9*). However, the precursors do not necessarily represent a final colibactin structure. Previous research suggested that colibactin alkylates DNA and forms a DNA adduct via a cyclopropane functional group, called the "warhead," which is structurally similar to other natural products that alkylate DNA (*9*, *10*). The importance of the cyclopropane ring was confirmed by identification of the colibactin resistance protein ClbS, which inactivates the cyclopropane ring to provide self-protection against DNA damage in the *pks*$^+$ bacteria (*11*). Recently, colibactin-DNA adducts with similar structures were detected in vitro using purified DNA and colibactin-producing bacteria (*10*). However, there was no direct evidence or structural characterization of these colibactin-DNA adducts in a biological setting.

Wilson *et al.* used an untargeted mass spectrometry DNA adductomics approach to structurally and mechanistically define a DNA alkylation end product of colibactin exposure. They identified two stereoisomeric colibactin adducts to the DNA nucleotide adenine in cultured mammalian cells and in colonic epithelial cells of formerly germ-free (sterile) mice colonized with a single *pks*$^+$ *E. coli* strain, providing direct evidence that these DNA adducts occur in vivo. As the authors note, the structure they uncovered does not necessarily represent the immediate colibactin-DNA adduct but is likely a degradation product of a larger colibactin adduct. This study provides important information about the structure and mechanism of action of colibactin. Furthermore, it describes a mass spectrometry method that could be used to identify other intractable compounds.

The adenine-colibactin adducts elucidated by Wilson *et al.* provide insight into how the cyclopropane functional group could react to alkylate DNA so effectively. These structures support a reaction mechanism whereby the cyclopropane ring is conjugated to an α,β-unsaturated imine formed from an intramolecular cyclodehydration that occurs once ClbP deacetylates the prodrug group. The presence of this proposed imine increases the reactivity of the cyclopropane ring to alkylate DNA

Model of colibactin-induced CRC

Precolibactins are synthesized from the *pks* island in *E. coli* before being activated by ClbP. When *E. coli* has direct contact with a mammalian cell, data suggest that the unstable, active colibactin reaches the nucleus where it alkylates DNA. A stable colibactin-DNA adduct was identified by Wilson *et al.*, revealing the structural identity of a biomarker for colibactin exposure and potentially for CRC risk.



GRAPHIC: JOSHUA BIRD/*SCIENCE*

(*12*) (see the figure). This alkylation generates DNA adducts that could lead to mutations in oncogenes or tumor suppressors that drive CRC tumorigenesis.

The identity of colibactin has been a long-standing question in the field of microbiota-influenced CRC. An important question to be resolved by further studies is how to distinguish the precise type of DNA damage responsible for the carcinogenic effects of colibactin. For example, what are the kinetics and relative levels of monoadducts versus interstrand DNA cross-links that can also result from alkylation and have been shown to occur after exposure to *pks*$^+$ *E. coli* (*6*, *13*)? Many other questions remain. For example, many bacterial biosynthetic gene clusters produce several bioactive molecules; is more than one colibactin variant produced from the *pks* island? Also, are there other roles for colibactin in mediating the interaction between the bacteria and human host? Undoubtedly, *E. coli* did not acquire *pks* to destroy its ecosystem by inducing DNA damage that may lead to cancer; instead, it is likely that *pks* imparts an important microbiological function, such as colonization and persistence in the gut (*14*).

From a clinical perspective, is there a way to predict which resident *E. coli* strains will colonize the gut mucosa and permit colibactin delivery? Colibactin requires direct cell-to-cell contact to exert its genotoxicity (*2*); thus, how does colibactin get from the bacteria into the nucleus of gastrointestinal epithelial cells, where it can cause DNA damage? Finally, how can we further apply our knowledge to improving clinical outcomes and treatment? This work has revealed a potential metabolite biomarker for colibactin exposure: adenine-colibactin adducts. However, it remains unknown whether adenine-colibactin adducts can distinguish precancerous tissue from healthy epithelium. We also do not yet know whether misrepaired adenine-colibactin adducts lead to gene mutations associated with known CRC subtypes and/or response to therapy. Future studies and the structural insight provided by Wilson *et al.* are expected to provide the next step toward applying microbiota signatures to improve prognosis and treatment for CRC. ■

REFERENCES

1. M. S. Donia *et al.*, *Cell* **158**, 1402 (2014).
2. J. P. Nougayrède *et al.*, *Science* **313**, 848 (2006).
3. G. Cuevas-Ramos *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11537 (2010).
4. J. C. Arthur *et al.*, *Science* **338**, 120 (2012).
5. C. M. Dejea *et al.*, *Science* **359**, 592 (2018).
6. M. R. Wilson *et al.*, *Science* **363**, eaar7785 (2019).
7. A. Cougnoux *et al.*, *Gut* **63**, 1932 (2014).
8. S. Tomkovich *et al.*, *Cancer Res.* **77**, 2620 (2017).
9. M. I. Vizcaino, J. M. Crawford, *Nat. Chem.* **7**, 411 (2015).
10. M. Xue *et al.*, *Biochemistry* **57**, 6391 (2018).
11. P. Tripathi *et al.*, *J. Am. Chem. Soc.* **139**, 17719 (2017).
12. A. R. Healy *et al.*, *J. Am. Chem. Soc.* **138**, 15563 (2016).
13. N. Bossuet-Greif *et al.*, *mBio* **9**, 1 (2018).
14. D. Payros *et al.*, *Gut Microbes* **5**, 313 (2014).

10.1126/science.aaw5475

STRUCTURAL BIOLOGY

Pathology-linked protease caught in action

Structural snapshots of γ-secretase yield insight for drug development

By **Stefan F. Lichtenthaler** *and* **Gökhan Güner**

The intramembrane protease γ-secretase has fundamental functions in animals, including signal transduction during embryogenesis and tissue homeostasis in adulthood. γ-Secretase cleaves its numerous substrates within their single transmembrane domains (TMDs), largely independently of their amino acid sequence. Abnormal cleavage of the substrates Notch and amyloid precursor protein (APP) is linked to leukemia and Alzheimer's disease (AD), respectively, making γ-secretase an important drug target for both diseases (*1*). Yet, chronic use of γ-secretase inhibitors (GSIs), such as in patients with AD, led to severe side effects, resulting from cleavage inhibition not only of the disease-relevant substrate APP but likely also of other substrates. Thus, there is a clear need to develop substrate-selective GSIs, but this requires a detailed understanding of how γ-secretase recognizes, binds, and cleaves its substrates. On page 708 of this issue, Zhou *et al.* (*2*) and another study by Yang *et al.* (*3*) provide a major step in this direction. Zhou *et al.* reveal the cryo–electron microscopy (cryo-EM) structure of human γ-secretase with its bound substrate, a fragment of APP. Yang *et al.* report a structure of γ-secretase, but bound with Notch. Together, the two studies demonstrate that binding of different substrates occurs in a similar manner and that both γ-secretase and substrate undergo specific structural rearrangements for substrate positioning in the active site. This has major implications for understanding the mechanism of γ-secretase and its function in signal transduction and AD, and for future development of substrate-specific GSIs with fewer side effects.

The aspartyl protease γ-secretase consists of four integral membrane protein subunits (*4*). The subunit presenilin (PS) contains the active site aspartyl residues (*5*) and exists in two variants, PS1 and PS2. Another subunit, nicastrin, has a tightly folded extracellular domain, which forms a lid on top of the membrane-bound γ-secretase complex. PEN-2 and APH-1A or APH-1B are additional subunits required for correct assembly, maturation, and trafficking of γ-secretase to the plasma membrane and endosomes.

Detailed biochemical analysis revealed that substrate cleavage by γ-secretase requires the substrate to move through an amazingly complex multistep process (see the figure). A substrate needs to have a short extracellular domain, either naturally (*6*) or as a result of an initial proteolytic cleavage (*7*), which is independent of γ-secretase and removes a large part of the extracellular domain, as for Notch and APP. This helps the substrate to fit below the lid imposed by nicastrin (*8*) and is considered a regulatory step to ensure that membrane proteins are only cleaved by γ-secretase when needed (*7*). Next, the truncated substrate likely binds to exosites outside of the active site of γ-secretase (*9*), followed by transfer to the active site, where cleavage occurs at the so-called ε site, a peptide bond close to the carboxyl-terminal end of the TMD in the substrate. Subsequently, the TMD is further truncated in a stepwise fashion up to the middle of the TMD (*10*), which is referred to as the final γ cut. If a membrane protein fails any of the requirements up to the ε cleavage, it will not be a substrate for γ-secretase.

Previously, a cryo-EM structure of γ-secretase was reported (*11*), but without the substrate, which was difficult to co-isolate. To achieve this, Zhou *et al.* and Yang *et al.* used two elegant tricks. First, they introduced single cysteine mutations into PS1 and either of the two substrates, derived from APP or Notch. The cysteines did not affect the activity of the protease or cleavage of the mutated substrate but allowed a stable cross-link between substrate and protease, essential for copurification. Second, one of the two catalytic aspartate residues in the active site was mutated to an alanine, which is known to abolish γ-secretase activity (*5*) and prevented undesired substrate cleavage during protein purification. Although both mutations are a caveat, the structures are in line with previous predictions based on bio-

German Center for Neurodegenerative Diseases (DZNE), Munich Cluster for Systems Neurology (SyNergy), School of Medicine, Klinikum rechts der Isar, DFG Research Unit FOR 2290 and Institute for Advanced Study, Technical University of Munich, Germany. Email: stefan.lichtenthaler@dzne.de

Revealing a microbial carcinogen

Rachel M. Bleich and Janelle C. Arthur

Science **363** (6428), 689-690.
DOI: 10.1126/science.aaw5475