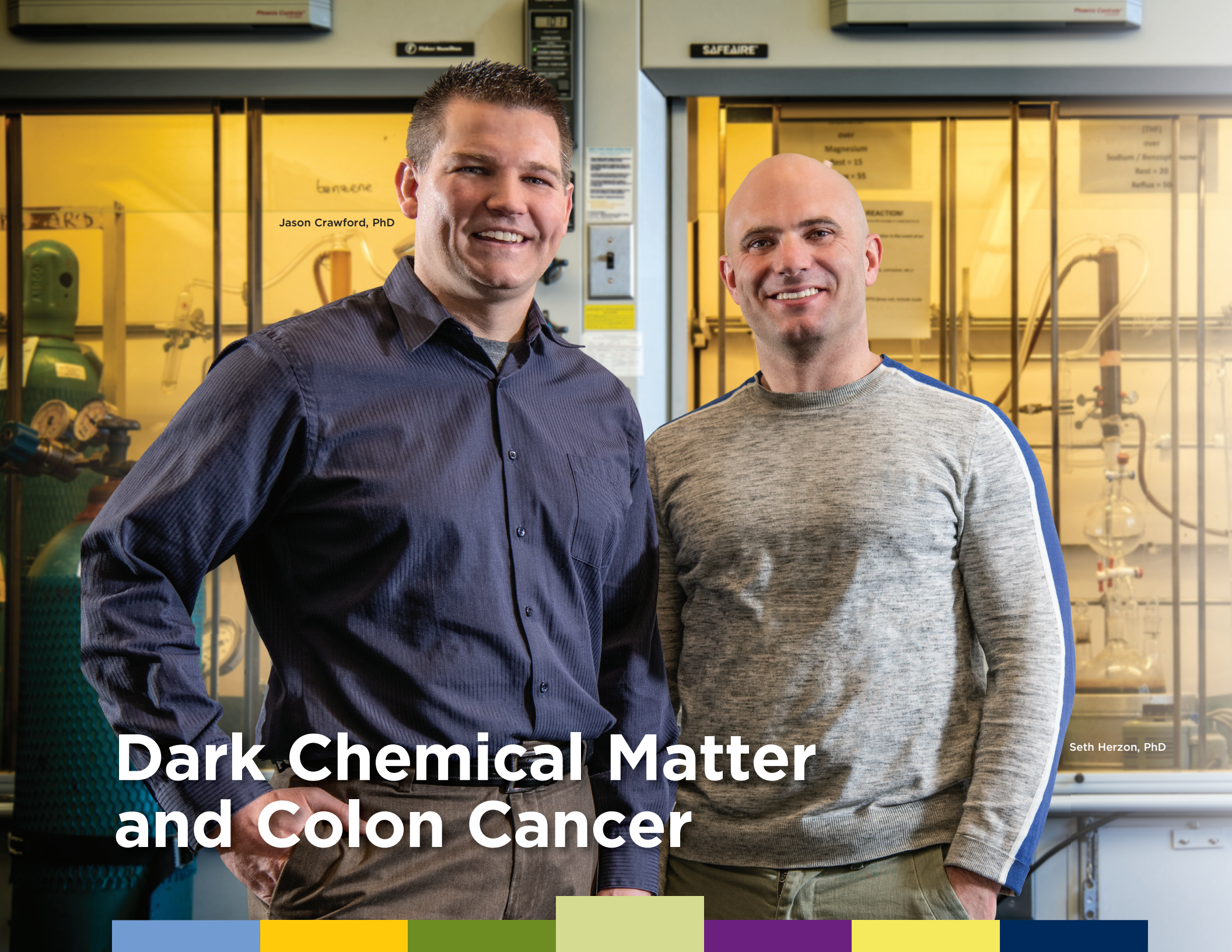
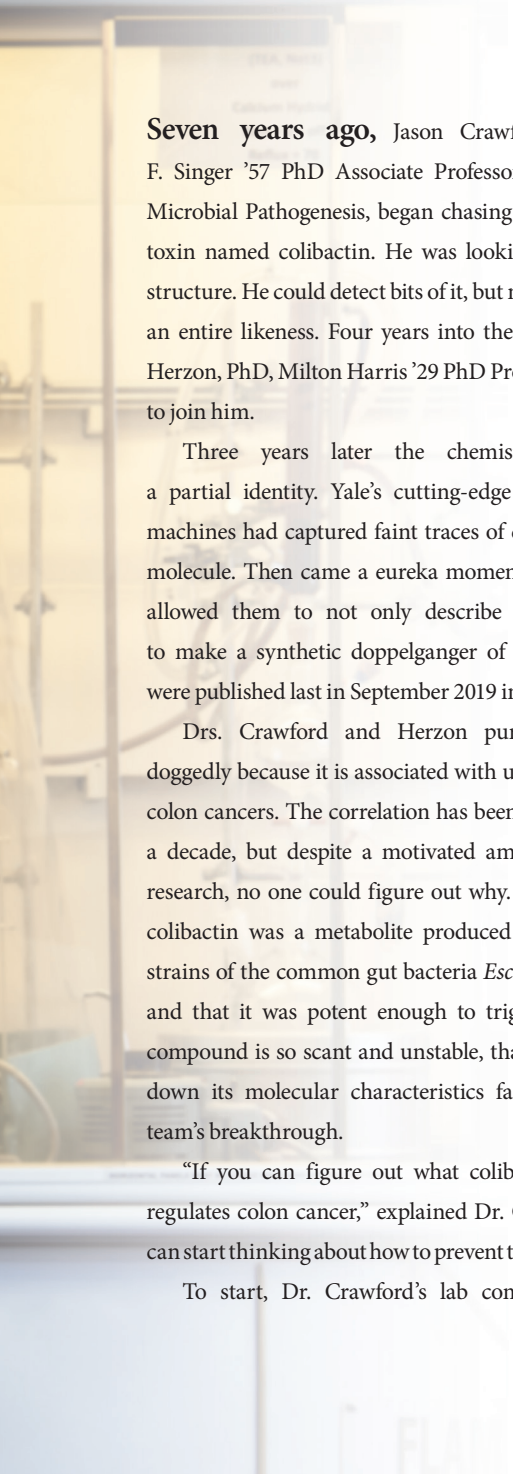


Dark Chemical Matter and Colon Cancer



Seven years ago, Jason Crawford, PhD, Maxine F. Singer '57 PhD Associate Professor of Chemistry and Microbial Pathogenesis, began chasing a ghost—a bacterial toxin named colibactin. He was looking for its molecular structure. He could detect bits of it, but never enough to form an entire likeness. Four years into the hunt he asked Seth Herzon, PhD, Milton Harris '29 PhD Professor of Chemistry, to join him.

Three years later the chemists still had only a partial identity. Yale's cutting-edge mass spectrometry machines had captured faint traces of colibactin's phantom molecule. Then came a eureka moment: the final clue that allowed them to not only describe colibactin, but also to make a synthetic doppelganger of it. Their discoveries were published last in September 2019 in the journal *Science.*

Drs. Crawford and Herzon pursued colibactin so doggedly because it is associated with up to 67 percent of all colon cancers. The correlation has been clear for more than a decade, but despite a motivated amount of competitive research, no one could figure out why. Scientists knew that colibactin was a metabolite produced by a few genotoxic strains of the common gut bacteria *Escherichia coli (E. coli),* and that it was potent enough to trigger cancer. But the compound is so scant and unstable, that all attempts to pin down its molecular characteristics failed—until the Yale team's breakthrough.

"If you can figure out what colibactin is and how it regulates colon cancer," explained Dr. Crawford, "then you can start thinking about how to prevent this from happening."

To start, Dr. Crawford's lab completed a metabolic analysis of ~40,000 molecules associated with the colibactin producer, eventually isolating and mapping more than 100 associated with the colibactin pathway. Then they stalled.

In 2015, Dr. Crawford mentioned to Dr. Herzon that he was working on a molecule that he couldn't isolate. Dr. Herzon, a synthetic chemist, said his lab probably could make a synthetic version in a couple of months. They began, as usual, with biomimetics, mimicking how nature might assemble the molecule. That process took them through "biosynthetic intermediaries" that revealed the chemical reactions that occur before colibactin becomes colibactin.

"We made the biosynthetic precursor, precolibactin, that some people had been studying and showed that that precursor transforms into a bunch of things that had been overlooked in the literature," said Dr. Herzon. "Building on the parts of colibactin that were well established, we began to make more complex fragments that we hypothesized might be generated. Those structures turned out to be extraordinarily potent genotoxins."

That excited the chemists, momentarily. Then they reached an impasse. They couldn't isolate the complete genotoxins from the bacteria. "By 2018," said Dr. Herzon, "I was thinking about making a graceful exit."

Then, in early 2018 another group of researchers cultured the bacteria that make colibactin, added DNA, and re-isolated the DNA. They found that the DNA now had an interstrand crosslink. This news reanimated the chemists. This synthetic compound, which Drs. Crawford and Herzon describe as containing two "warheads," was not colibactin itself, but behaved like the natural molecule. They hypothesized that colibactin itself was stuck in the crosslink.

"We decided to use the DNA as a hook to fish out colibactin," explained Dr. Crawford.

Dr. Herzon asked one of his graduate students, Mengzhao (Lucy) Xue, to see if she could dig the molecule out of the crosslink and isolate it. Xue crosslinked DNA with the bacteria, then used enzymes to gnaw down the DNA until all that remained were the two DNA bases and the molecule that had reacted with them. Using extremely powerful mass spectrometry and isotope labeling, Xue eventually found the molecule's mass.

"That was the bit of data we needed," said Dr. Herzon. "It was the most exciting scientific moment of my career." This data allowed them to predict a structure for colibactin, and then chemically synthesize it.

"That was a magic moment," agreed Dr. Crawford. "It took seven years to get there. We quickly went from not having a clue to knowing almost everything about the entire mechanism, because we could mine seven years of data."

Dr. Crawford is developing an antibiotic that selectively removes colibactin-producing bacteria from the intestinal tract, with the goal of preventing colon cancer. Dr. Herzon hopes to repurpose colibactin as targeted chemotherapy.

"There is probably a world of metabolites out there that are biologically significant and relevant to human health but that pass completely undetected by existing methods. We refer to it as, 'Dark Chemical Matter.' You can't detect it, but it's there and it probably matters a lot." Looking forward, the researchers are trying to develop approaches to reveal the structures and function of this dark chemical matter.