A Conversation with Emmanuelle Charpentier

INTERVIEWER: GUY RIDDIHOUGH

Senior Editor, Science

Emmanuelle Charpentier is Director of the Department of Regulation in Infection Biology at the Max Planck Institute for Infection Biology in Berlin and a Visiting Professor at the Laboratory for Molecular Infection Medicine Sweden at Umeå University’s Centre for Microbial Research.

Guy Riddihough: There’s a system of repeats in bacteria that are important in genome engineering and potentially for human therapeutic purposes. Can you set the stage for us with regard to these repeats?

Dr. Charpentier: A Japanese group described a series of repeats in bacteria in the 1980s. This was not followed up until the mid-to-late 1990s, when some researchers found out that these repeats could be transcribed as RNAs that could be processed. Begin in 2000 and into the mid-2000s, some bioinformaticians figured out that next to the repeats, you had a set of genes encoding the so-called CRISPR (clustered regularly interspaced short palindromic repeats)-associated proteins and that those proteins had characteristics of DNA- and RNA-interacting proteins. The idea was that the system might work as an adaptive immune system that would allow bacteria and archaea to defend themselves against mobile genetic elements, such as plasmids or phages.

Guy Riddihough: What was the evidence that it was a defense system?

Dr. Charpentier: There were two lines of evidence. First, the genes encoding proteins with putative DNA- and RNA-interacting functions. Then there was the fact that there were short sequences interspersed between the repeats called spacer sequences. Those spacer sequences originate from mobile genetic elements, as short pieces of sequences from these genomes. The idea was that these very small sequences were actually memory devices of past infections. The first time bacteria or archaea encounter the mobile genetic element like a phage or a plasmid, the CRISPR–Cas system recognizes the entering DNA of this invading foreign mobile genetic element and cleaves a piece of DNA from it to insert into the CRISPR array, allowing the bacteria or archaea to “remember” the infection. Then, upon a second infection, this memory sequence is transcribed as CRISPR RNA that will guide a series of CRISPR-associated (Cas) proteins to the invading genome. The system thus deals with infections by using homologous base-pairing between the CRISPR RNA that contains the memorized mobile genetic element and the DNA or RNA of the cognate invading element.

Guy Riddihough: Your entry into this field was through small RNAs.

Dr. Charpentier: Yes. In my lab, we focus on human bacterial pathogens; we are mainly working with a human pathogen called Streptococcus pyogenes, also known as Group A Streptococcus. We are interested in how different regulatory mechanisms of gene and protein expression coordinate to allow this bacterium to adapt to and survive in its environment and cause diseases in humans, as well as how the human host defends itself against the infection by S. pyogenes.

Around 2006, we performed bioinformatics screens searching for novel small regulatory RNAs (sRNAs) that are not supposed to code for proteins (noncoding RNAs). We found a lot of them, but we focused on one that is now known as tracrRNA (trans-activating crRNA). We also identified the fact that this tracrRNA was located next to the Type II CRISPR–Cas system.

Guy Riddihough: Why did you pick that particular RNA?

Dr. Charpentier: We were validating the expression of putative sRNAs that we identified using northern blot analysis. tracrRNA was very well expressed in northern blots. We had mapped this RNA and found that it was well expressed as two primary transcripts and one processed form. I just had a gut feeling that it could work in trans, compared to the other classes of RNAs that work in cis or that belong to the family of riboswitches. We also used bioinformatics tools to determine if this regulatory RNA could target a messenger RNA. We found a target messenger RNA encoding a virulence factor that can base pair with tracrRNA. However, we could not make sense of this interaction. We were already interested in working on the CRISPR–Cas system in S. pyogenes because small RNAs are involved and...
CRISPR RNA is just a small RNA. In *S. pyogenes*, the spacers target temperate phages that carry genes that contribute to the virulence of the bacterium. For us, it was a way to show for the first time that CRISPR–Cas could have a role in virulence by limiting the acquisition of virulence genes. At some point, because it was clear that tracrRNA is located in the vicinity of the type II CRISPR–Cas9 system, I concluded that tracrRNA and the CRISPR system might have something to do with each other.

When we first started working on CRISPR–Cas, Sylvain Moineau, Philippe Horvath, and Rodolphe Barrangou had just shown that the system was indeed an adaptive immune system against lytic phages that kill bacteria, so we thought that, theoretically speaking, it would be nice to have another look at the function of the system. So we started to test this hypothesis using genetics and by developing read-out systems of temperate phage infections. Elizita Deltcheva—a master’s student in my lab—and Krzysztof Chylinski looked at knockouts of both tracrRNA and CRISPR RNA that we had generated using precise gene surgery tools we had developed in the lab to see if the knockout of each would change the expression of the other. Indeed, the two RNAs needed each other to coprocess one another. This is how we started working on CRISPR–Cas in parallel with tracrRNA.

**Guy Riddihough**: The tracrRNA is a distinguishing feature compared to the Type I and Type III CRISPR–Cas systems. Can you explain briefly and simply how the type II system works?

**Dr. Charpentier**: tracrRNA contains an antirepeat that allows it to base pair with each of the repeats of the CRISPR array, forming a duplex RNA. The exact mechanism by which the two RNAs meet in the bacterial cell isn’t understood especially well, but in vivo Cas9 (formerly named Csn1) stabilizes the duplex. RNase III, which is an enzyme produced by the bacterium, recognizes this RNA duplex and cleaves it. This gives rise to a complex that includes the Cas9 enzyme, an intermediate form of the CRISPR RNA containing a portion of the repeat and the full spacer portion that targets the phage genome, and another part of the repeat that’s bound to the tracrRNA.

We’re still trying to decipher this pathway. There was the possibility that the tracrRNA might somehow leave the system at some point if it was only important to help the maturation of the CRISPR RNA. Instead, my laboratory thought and found that the mature form of the complex maintains the RNA duplex throughout, stabilized by Cas9. Then, both RNAs guide Cas9 to the invading genome, where Cas9 uses two nuclease domains specific to that enzyme to cleave each strand of the double-stranded DNA. tracrRNA’s interaction with the CRISPR RNA is an integral part of the Cas9 guidance system.

**Guy Riddihough**: At what point did you realize that this is interesting not only in terms of understanding how prokaryotes protect themselves against invading nucleic acids, but also has huge biotechnological applications?

**Dr. Charpentier**: I’m a bit obsessed to find ways to exploit what we do in the lab for new ways to treat diseases. I started in the field of medical microbiology and we work with bacterial pathogens, so we always try to find interesting pathways and mechanisms that could lead to new strategies for therapeutics. In the past, I developed genetic tools to facilitate the manipulation of Gram-positive bacteria and I’ve also worked with mouse models. When you work on small RNAs with antisense mechanisms, you always know that in principle you might find a new genome-silencing mechanism that could be exploited further, like RNAi.

What was new, and specific, to CRISPR–Cas is that it targets DNA. We knew that small RNAs in bacteria can target proteins and titer them away from their function, or that they can interact with and affect the stability or the translation of messenger RNAs. What was missing was something like CRISPR–Cas, which is an RNA-guided system that targets DNA.

I could see that this tool could be used for the genetics not only in bacteria, but also for genetics in eukaryotes, that it could be exploited for the treatment of human genetic disorders. TALENs (transcription activator-like effector nucleases) and zinc finger nucleases were described a couple of years ago and were already recognized as precise gene surgery tools that functioned well in eukaryotes. People switched to CRISPR–Cas9 because in terms of identifying the target to be cleaved by the system, it’s easier to direct and hybridize your nuclease using a programmable RNA than to have to design, for example, a much more complicated zinc finger nuclease.

**Guy Riddihough**: Since your work’s been published there’s been a complete revolution in genome engineering. Could you tell us about that, and some of the other applications that CRISPR–Cas9 has been used for?

**Dr. Charpentier**: Biologists really needed a tool that would allow them to perform easily precise surgery in the genomes of eukaryotes. It can be really frustrating to not be able to manipulate the genome as you would like, as you could in bacteria. CRISPR–Cas facilitates our understanding of gene functions because it allows biologists to perform knockouts easier and to correct mutations more precisely. It helps biologists working in epigenetics because it has been developed to modify genes at will. It can also be harnessed to inhibit or activate gene expression at the level of the RNA, to down-regulate or up-regulate the expression of genes.

It’s obviously very popular now as a tool to engineer humanized mouse models of diseases. It’s interesting to be able to study the causes of diseases but also to be able to develop screening methods to find new targets for therapeutics and to use these models to test drugs in development. It has a real impact in academics but also in the biotechnology and the biopharmaceutical industry.
Guy Riddihough: It seems to work so well in so many different situations, in different organisms, different cells. Obviously, you don’t have to introduce many components to make the system usable.

Dr. Charpentier: That’s another reason it has been quickly and largely adopted by the community. The tools are available to express the CRISPR–Cas system and bring it to the nucleus, which does not exist in bacteria. You just need to express the Cas9 enzyme and the specific single-guide RNA using whatever vectors or delivery systems are already in use in your lab, and the system will work.

Guy Riddihough: One of the ultimate goals is treatment of human diseases, but that’s not without its technical issues as well.

Dr. Charpentier: Human gene therapy has made a lot of progress lately. Again, scientists really needed a tool that facilitated the precise manipulation of genes. Depending on the cell system and depending on the way CRISPR–Cas9 is expressed and delivered, it can be very specific. I’m sure that in the years to come we’ll see developments with regard to the technology and its delivery to treat human diseases, probably by combining gene therapy with cell therapy.

Guy Riddihough: Tell us a bit about other areas of research in your lab.

Dr. Charpentier: We are really interested in RNA- and protein-mediated regulation. In addition to the small regulatory RNAs, we work on the RNases that are important to their biology, and also to the degradation of messenger RNA. We also focus on understanding the control of protein quality and turnover in bacteria. When bacteria encounter stress conditions, they can accumulate misfolded proteins or aggregates of proteins that need to be resolved by specific proteolytic machineries. Otherwise, those aggregates or misfolded proteins could be toxic for the cells. We’ve focused on a family of proteins called Clp proteins that are involved in so-called general proteolysis, the proteolysis that is triggered when bacteria encounter stress conditions. We also work on the regulation of proteolysis by specific factors during the normal life of the bacteria or certain adaptational processes. We’re always looking at those mechanisms within the context of the interactions between the bacteria and the host. We have other projects regarding cell–cell communication between the bacteria and the host and the host response to bacterial infection.