Title: Structure guided characterization of bacterial antiviral defense systems

PI: Dinshaw J. Patel, Memorial Sloan-Kettering Cancer Center, New York, NY

Funding: NIH GM129430, AI141507 and GM145888

Abstract

The arms race between bacteria and the most abundant organisms in the biosphere, namely bacteriophages, gave rise to the dynamic bacterial pan-genome composed of a combination of sensor-effector systems that were lost and gained, thereby evolving with, and combating against, the incessant phage insults. Numerous defense strategies employed by bacteria, including CRISPR-Cas and CBASS (cyclic oligonucleotide based antiphage signaling systems) surveillance complexes, associate with each other in pan-genomic defense islands, thereby providing defense against a variety of phages (Berheim et al., 2020; Tesson et al., 2022). Recently, additional antiphage defense systems have been identified, named after the protective gods and goddesses of various mythologies, often include pre-existing genes from the bacterial genetic repertoire that have undergone key modifications to acclimatize with the phage-rich environment. We propose to apply cryo-EM methodology to structurally characterize the membrane-spanning Cam1, as well as Lamassu, Zorya and Azaca antiphage defense systems comprising of a diverse set of sensor and effector modules that execute a common altruistic cell killing phenomenon, termed abortive infection, thereby enabling bacterial cell death prior to phage-induced lysis. While a few of these systems including Thoeris and Gabija have been characterized recently (Ofir et al., 2021; Duncan-Lowey et al., 2023; Sadie et al., 2023), the underlying mechanism behind a vast majority remains elusive. To this end, structure-function studies could revolutionize our understanding of antiphage defense pathways with biomedical applications.

Proposal:

The bacterial defense related genes are found to be spatially co-localized in bacterial genomes and are shared amongst a population of bacteria by horizontal gene transfer, which reduces the energy burden from a single bacterium (Bernheim & Sorek, 2020; Makarova et al., 2011). Some of these systems involve different components, function in a synchronized fashion upon viral infection, whereas others form large complexes (Duncan-Lowey et al., 2023). In either case, the defense systems have the potential for viral infection recognition, and in some cases this step is enough to activate the effector molecule, which kills the infected bacteria and aborts further propagation of the virus. In some systems there are mediators, also referred to as cyclic oligonucleotide secondary messengers, which are generated by bacterial enzymes like cyclases and relay the signal to the effector molecules upon viral infection (Cohen et al., 2019; Whiteley et al., 2019; Patel et al. 2022). The mechanism by which these enzymes are activated to generate the second messenger in the virus-infected bacterial cells remains unclear.

Moreover, the activated effector molecules follow different approaches to abortive infection – the most common effectors are nucleases, which degrade the viral genetic material selectively or without any self-discrimination; this family of effectors are relatively well characterized. However, there are other effectors which incorporate a transmembrane domain, that function either by membrane depolarization, metabolite depletion by pore formation or through disintegration of the bacterial cell membrane (Duncan-Lowey et al., 2021; VanderWal et al., 2023). To date, we have very limited understanding about the mechanism of the membrane-associated effectors.

We outline several proposed Aims below:

Aim 1: Structure-guided mechanism of transmembrane effector protein Cam1 activation (collaboration with the **Luciano Marraffini** lab at Rockefeller University)**.**

Experienced cryo-EM users: Drs. Puja Majumder and You Yu

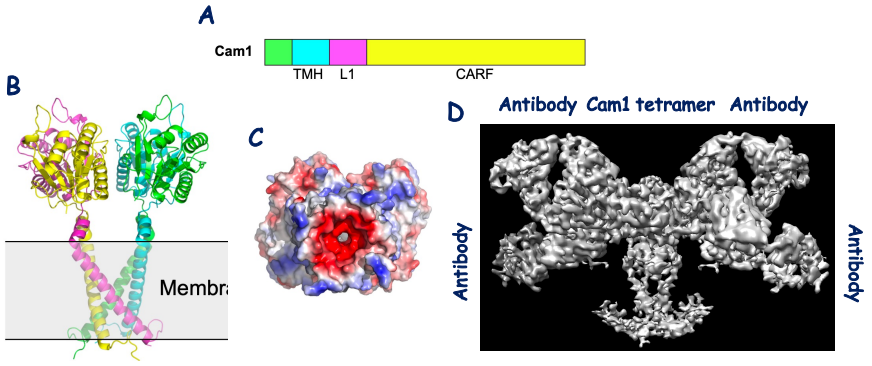

Figure 1. (A) Domain architecture of Cam1. (B) AlphaFold-II model of full-length Cam1 tetramer. (C) Surface electrostatic view looking down the channel of AlphaFold-II model of full-length Cam1 tetramer. (D) Electron density map of 3.9 Å cryo-EM structure of ligand-free full-length Cam1 tetramer.

Type III CRISPR-Cas immune systems in bacteria provide adaptive immunity against viruses through usage of CARF domain effector proteins. Though most of the CARF domain effectors which function as nucleases are relatively well characterized, the membrane-associated CARF domain effectors remain poorly understood (Shah et al., 2019; Shmakov et al., 2018). Here, our efforts are focused on understanding the abortive infection mechanism of a transmembrane domain containing

CARF-effector protein, Cam1, composed of a single transmembrane helix connected by linker L1 to a CARF domain (**Fig. 1A**). The Marraffini lab has used functional approaches to establish that Cam1 mediates membrane depolarization to provide phage defense during the type III CRISPR-Cas immune response. Our lab has shown that the CARF domain of Cam1 dimer binds cA_4 second messenger and Dr. Yu has solved x-ray structures of CARF domain-only Cam1 dimer in the apo- and cA_4-bound states. Further, using SEC-MALS, Dr. Majumder established that full-length Cam1 forms a dimer-of-dimers (tetrameric) topology in GDN (glyco-diosgenin) detergent, with AlpfaFold-II predicting formation of a tetrameric positively-charged transmembrane pore (**Fig. 1B, C**). Dr. Majumder has generated antibodies targeted to the CARF domain-only and subsequently formed full-length Cam1-antibody complexes in the absence and presence of cA_4 in GDN (glyco-diosgenin) detergent. We have collected cryo-EM data on the ligand free full-length Cam1-antiobody complex in GDN and have processed the data set to 3.9 Å. The electron density map is shown in **Fig. 1C** establishing formation of a full-length Cam1 tetrameric complex stabilized by four antibodies with the channel formed by four transmembrane helices. Currently, she is switching particle selection from template picking to the TOPAZ peak picking program towards improving the resolution.

Her next goal is to collect data on the cA_4-bound full-length Cam1-antibody complex in GDN and solve its structure. A comparison of the free and cA_4-bound Cam1 structures should provide structural insights into conformational transitions that contribute to the triggering of abortive infection on complex formation. Notably, our structural studies definitively establish apparently for the first time that a transmembrane pore can be generated by alignment of four alpha helices.

Aim 2: Lamassu antiphage defense system with nuclease and hydrolase/protease as effectors (collaboration with the **Samuel Sternberg** lab at Columbia University)**.**

Experienced cryo-EM users: Drs. Arpita Chakravarti and You Yu

The Lamassu antiphage defense system, which has pre-established roles in mitosis and meiosis (Harvey et al., 2002, Hirano et al., 2005), is composed of a Structural Maintenance of Chromosome (SMC) sensor (LmuB) and partner Kleisin subunit (LmuC) that is predicted to sense phage DNA, a variable effector (LmuA) with diverse cell-killing functions and an unknown C-terminal domain, that upon activation, ensues an altruistic cell death mechanism known as abortive infection, together with an unknown component that may or may not be present in the operon (Nasmyth et al., 2005). Since Lamassu exists in variable operonic compositions owing to variable NTDs and number of effectors, we decided to begin our studies with one possessing a Cap4 nuclease as a LmuA effector (**Fig. 2A, top**). In this regard, we are quite interested in understanding how antiphage SMCs evolved and deviated from their established roles in cell cycle and chromosomal architecture (examples include DNA looping by condensin, cohesion and Smc5/6 systems), to DNA repair (Mre11-Rad50 complex) (Wu et al., 2012), and eventually as a sensor in Wadjet (Liu et al., 2022; Deep et al., 2022) and Lamassu systems involved in biological conflicts.

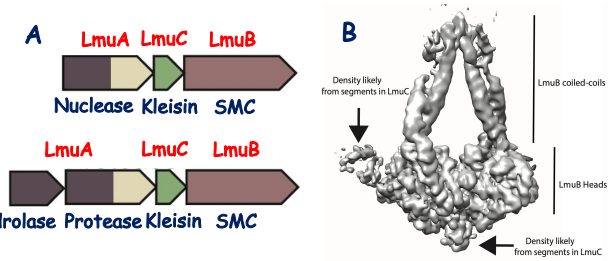

Figure 2. (A) Lamassu systems containing Cap4 nuclease effector (top panel) and hydrolase-protease effector (bottom panel). (B) Electron density representation of the 3.3 Å cryo-EM structure of LmuBC binary complex.

From a structural perspective, cryo-EM would be a powerful tool to obtain high resolution structures of all the Lamassu components separately and as complexes together with nucleic acids, which could be potential activators of the system. Additionally, structural insights into the unknown CTD of the effector would be key in terms of understanding protein-protein interactions in this system since it has been predicted to sequester its partner proteins. Our group has extensive ongoing experience on cryo-EM studies of the DNA-bound SMC Smc5/6 system (Ye et al. 2022) involved in DNA repair and recombination, as well as the DNA-bound Mre11-Rad50 system involved in DNA double strand break repair (Hol et al. in preparation), that should help us move forward on the structural characterization of the Lamassu system.

To this end, Dr. Chakravarti has expressed and purified *Vibrio cholerae El Tor* biotype LmuBC binary, the LmuABC ternary and the LmuABC-DNA quarternary complexes. This required considerable biochemical effort to overcome aggregation issues. Cryo-EM data has been collected and processed on the LmuBC binary complex on quantifoil grids, with the 3.3 Å structure in density representation shown in **Fig. 2B**. LmuB exhibits the characteristic dimeric SMC fold composed of head and head-proximal coiled-coil domains, while short segments of the LmuC Kleisin domain are visible in the map. Dr. Chakravarti has had to shift to graphene oxide grids to collect cryo-EM data on the LmuABC ternary complex and this data is being currently processed in efforts to define the position of the LmuA effector nuclease relative to the LmuBC SMC component and should also provide insights into the oligomeric state of the complex. The ultimate goal of the project is to solve the cryo-EM structure of the LmuABC-DNA quarternary complex (containing a LmuA nuclease dead mutant) to define whether DNA is encapsulated within the SMC complex and also define its positioning relative to the nuclease pocket. A notable feature of the Lamassu system is that LmuB

SMC subunit has a substitution in the Walker B motif that prevents hydrolysis of ATP, unlike SMC proteins involved in DNA repair and recombination that can hydrolyze ATP.

Furthermore, a long-standing but poorly understood question in the field relates to the origin and identification of triggers that activate the sensor modules in antiphage defense systems. To this end, we propose to undertake structural studies of Lamassu components with a variety of different nucleic acid substrates, that could serve as potential activators, towards deciphering the underlying mechanism of activation. In addition to focusing on *in vitro* structural and biochemical aspects of the Lamassu defense system towards elucidation of key protein-protein and protein-DNA interactions, such efforts will be followed by *in-vivo* phage-based functional studies (in collaboration with the Samuel Sternberg lab) that would potentially uncover the mechanism of this elusive system, thereby explaining why house-keeping genes were co-opted for secondary functions in the bacterial pan-genome and how they execute this distinct role with the help of their partners.

In addition to Cap4 nuclease effectors, Lamassu possesses effectors that execute hydrolase and protease activities to ensue abortive infection (**Fig. 2A, bottom**). This system has turned out to be more tractable biochemically and Dr. Chakravarti aims to generate LmuABC complexes and perform cryo-EM studies that should potentially provide an understanding of the need for two effectors in this system along with their interaction with its sensor SMC. Further, a comparison of the hydrolase/protease Lamassu system with its with Cap4 nuclease counterpart, would likely highlight similarities or differences in mechanisms of different Lamassu systems.

Aim 3: Uncovering the mechanism of the Zorya and Azaca antiphage defense systems

Experienced cryo-EM users: Dr. Puja Majumder
New Cryo-EM user: Dr. Abhishek Suman

Our longer-term goal is to expand our ongoing research on the Lamassu antiphage system to its counterpart Zorya and Azaca antiphage systems. Our interest in deducing strategies of antiphage pathways led us towards the intriguing Zorya system which is unique owing to its unusual operonic composition of two proteins (ZorA and ZorB) that are predicted effectors likely to form membrane pore complexes akin to bacterial flagellar stator complex (MotA and MotB) involved in proton transport (Santiveri et al., 2020) and two other components (ZorC and ZorD), one of which (the former) is a domain of unknown function while the other (the latter) being a SNF2 helicase, which peculiarly, could be the likely sensor (**Fig. 3A**) (Doron et al., 2018). Structural characterization of Zor subunits individually and in complex with others would help deducing the mechanism underlying antiphage defense of this system. This project is in its infancy, and to date, Dr. Suman is currently expressing ZorC and ZorD, while Dr. Majumder is using her experience in membrane proteins in attempting to express and purify the membrane-spanning subunits ZorA and ZorB. The plan is to investigate by cryo-EM the structures of various combinations of Zor subunits towards an understanding of its subunit architecture and the principles underlying its role in abortive infection.

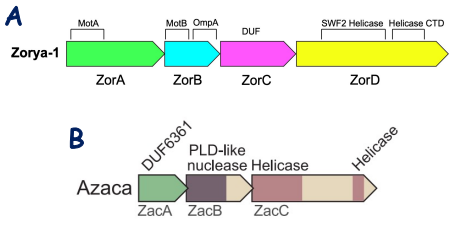

Figure 3. Domain architecture of (A) Zorya-I and (B) Azaca antiphage defense systems.

The Azaca antiphage defense system is composed of ZacC, a domain of unknown function, ZacB, a PLD-like nuclease and ZacC, a helicase (**Fig. 3B**). To date, nothing is known about Azaca's role in antiphage defense, with Dr. Majumder preparing constructs for generation and purification of individual subunits, prior to their individual and multi-subunit cryo-EM-based structural characterization.

References

Bernheim, A., & Sorek, R. (2020). The pan-immune system of bacteria: antiviral defence as a community resource. *Nat Rev Microbiol*, *18*(2), 113-119.

Cohen, D., et al. (2019). Cyclic GMP-AMP signalling protects bacteria against viral infection. *Nature*, *574*, 691-695.

Deep, A. et al. (2022). The SMC-family Wadjet complex protects bacteria from plasmid transformation by recognition and cleavage of closed-circular DNA. *Mol. Cell* 82, 4145-4159..

Doron, S., et al. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* 359:6379.

Duncan-Lowey, B., et al. (2021). Effector-mediated membrane disruption controls cell death in CBASS antiphage defense. *Mol Cell* 81, 5039-5051.

Duncan-Lowey, B., et al. (2023). Cryo-EM structure of the RADAR supramolecular anti-phage defense complex. *Cell*, *186*, 987-998.

Harvey, S. H., Krien, M. J., & O'Connell, M. J. (2002). Structural maintenance of chromosomes (SMC) proteins, a family of conserved ATPases. *Genome biology*, 3, REVIEWS 3003.

Hirano T. (2005). SMC proteins and chromosome mechanics: from bacteria to humans. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 360, 507–514.

Liu, H. W. et al. (2022). DNA-measuring Wadjet SMC ATPases restrict smaller circular plasmids by DNA cleavage. *Mol. Cell* 82, 4727-4740.

Makarova, K. S., Wolf, Y. I., Snir, S., & Koonin, E. V. (2011). Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *J Bacteriol*, 193, 6039-6056.

Millman, A. et al. (2022). An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host & Microbe* 30, 1556-1569.

Nasmyth, K., & Haering, C. H. (2005). The structure and function of SMC and kleisin complexes. *Annual review of biochemistry*, 74, 595–648.

Ofir, G. et al. (2021). Antiviral activity of bacterial TIR domains via immune signalling molecules. *Nature*, 600, 116–120.

Patel, D. J., Yu, Y. and Jia, N. (2022). Bacterial origins of cyclic nucleotide-activated antiviral immune signaling. *Mol. Cell* 82, 4591-4610.

Peng X. P. & Zhao, X. The multi-functional Smc5/6 complex in genome protection and disease. *Nat. Struct. Mol. Biol.* 30, 724-734.

Sadie P. et al. (2023) Structural basis of Gabija anti-phage defense and viral immune evasion. bioRxiv 2023.05.01.538945.

Santiveri, M. et al. (2020). Structure and function of stator units of the bacterial flagellar motor. *Cell* 183, 244-357.

Shah, S. A., et al. (2019). Comprehensive search for accessory proteins encoded with archaeal and bacterial type III CRISPR-cas gene cassettes reveals 39 new cas gene families. *RNA Biol*, 16, 530-542.

Shmakov, S. A. et al. (2018). Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene neighborhood analysis. *Proc Natl Acad Sci U S A*, 115, E5307-E5316.

VanderWal, A. R., et al. (2023). Csx28 is a membrane pore that enhances CRISPR-Cas13b-dependent antiphage defense. *Science*, 380, 410-415.

Tesson, F., et al. (2022). Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat Commun* 13,: 2561.

Whiteley, A. T., et al. (2019). Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature*, 567 194-199.