Deciphering Global Dynamics of Microbial Multidrug Resistance Regulation using Novel method CLK-seq

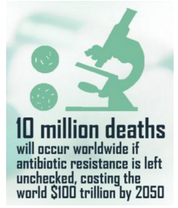
Kunal Poorey*[1], Aashish Priye[2], Deanna J Curtis[1], Annette E. LaBauve[2], Ronald F. Renzi[3], Meiye Wu[2,4] & Robert J. Meagher[2]

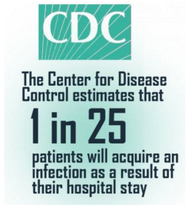
1: Systems Biology, 2: Biotechnology and Bioengineering, 3: Advanced Systems Engineering & Deployment, 4: Current: BioRad, Hercules, CA

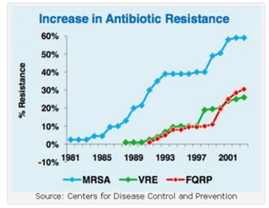
SAND2018-0114C

Abstract

The emergence of multi-drug resistant (MDR) bacteria translates to $20 billion per year in healthcare expenses in the US alone. Furthermore, the lack of tools to study these emerging pathogens has made it difficult to devise effective treatments. If left unaddressed, MDR bacteria will pose an enormous threat to US national security by crippling both the economy and the health of the population. For this project, we conducted a study aimed towards understanding MDR bacterial genetic responses to antibiotics, to elucidate the molecular mechanisms underlying the regulation and evolution of antibiotic resistance.





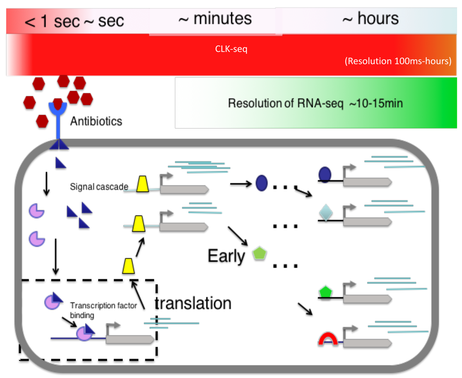


Current genomic methods do not have the temporal sensitivity to analyze early transcriptional responses of bacteria to antibiotic challenges. To address this gap, we have developed a novel assay called Crosslinking Kinetic Sequencing (CLK-Seq) to better understand resistance mechanisms and identify new biomarkers for drug resistance in MDR bacteria. CLK-Seq is a novel universal assay to characterize the dynamics of DNA/protein interactions of any organism without the use of antibodies, which was used to characterize an MDR pathogen *K. pneumoniae* for the first time. We combine CLK-seq study with RNA-seq studies to characterize the gene regulation of this pathogenic bacteria to understand the underlying principles of MDR at the molecular level.

We measured gene expression profiles, using RNA-seq, of these MDR bacteria in steady state and in presence five classes of antimicrobial drugs, clinically relevant for Gram-negative bacterial infections. We compare these responses to the untreated samples to see how these pathogens function to survive such harsh treatments. We find that pathogens response to different drugs varies and both in types of genes and also in strength. We further took the samples which showed the greatest gene expression response to an antibiotic drug to apply CLK-seq assay to characterize the gene regulation. We narrow down a few genes which keystone markers for the cascade of gene expression changes. We also find evidence of changes in major structures in DNA compaction correlated with gene transcriptional changes. The CLK system has enabled the collection of kinetic information of protein/DNA interactions to reveal underlying mechanisms of MDR regulation and its dynamics at a time scale of seconds.

This holistic multi-omics approach information is crucial to understand the underlying mechanisms and pathways for development of multi-drug resistance, their functions in infection and pathogen survival. This information arms us with new candidate biomarkers for more effective drugs against these emerging superbugs.

Background



Early Regulators

A schematic of a hypothetical scenario of a multidrug-resistant pathogenic bacterial response to a drug. When drug interacts with the pathogen, the quorum sensing receptors sense the toxins/drugs and activates a signaling cascade to activate the first cascade of gene expression change. These mRNAs will translate to proteins which will either nullify the effect of the drugs and/or regulate the next wave of genes further and this goes on as long as the pathogen is stressed. The top bar of the figure shows the times line of the events from seconds to hours. The time resolution of RNA-Seq assay is about 10-15 minutes at best, but time resolution for the CLK-seq in previous studies in yeast have suggested can be seconds. CLK-seq can provide more time-resolved useful information than RNA-seq to decipher and understand the ABR mechanism including crucial early regulators.

We have developed a mathematical model for the DNA binding and crosslinking kinetics. The illustration explains the boundary conditions and the final model, and the panel on right shows the behavior of derived model as a function of crosslinking time. This model when applied to experimental data, derives kinetic parameters for protein DNA interaction of a relevant site.

$$\underset{unbound\ TF}{\theta_u} \underset{r_{1b}}{\overset{r_{1f}}{\rightleftarrows}} \underset{bound\ TF}{\theta_b} \xrightarrow{r_2} \underset{crosslinked\ TF}{\theta_{xl}}$$

$$\frac{d\theta_b(t)}{dt} = r_{1f} - r_{1b} - r_2\,,\ \frac{d\theta_{xl}(t)}{dt} = r_2 \quad \text{where}$$

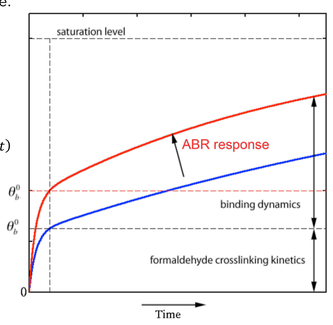
$$r_{1f} = k_a C_{TF}\theta_u(t), \qquad r_{1b} = k_d\theta_b(t) \quad \&\quad r_2 = k_{xl}C_{FH}\theta_b(t)$$

Boundary Conditions
At t = 0

$$\theta_{xl}(0) = 0, \theta_b^0 = \frac{k_a C_{TF}}{k_a C_{TF} + k_d}$$

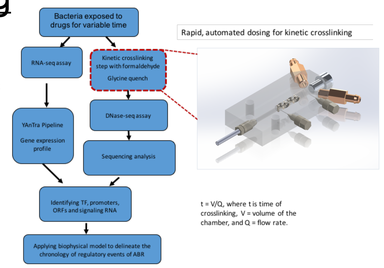
With assumptions $(k_a C_{TF} + k_d)/k_{xl}C_{FH} \ll 1$

$$\theta_{xl}(t) \approx 1 - (1-\theta_b^0)e^{-k_a C_{TF} t}$$

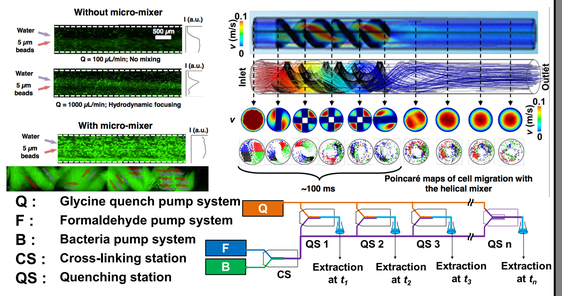


Strategy and Engineering

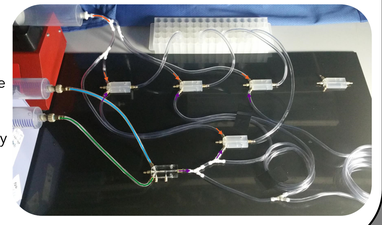
For this study, we concentrate our effort to demonstrate the application of CLK-seq assay on MDR NDM1 containing *K. pneumoniae*.
The pathogen is exposed to sublethal dose of antibiotics for a variable amount of time and studied by two different methods.
Transcriptional response to the drug treatment is studied with RNA-seq and, CLK-seq/DNase assay is applied to measure the dynamics and binding activity of the regulatory elements in the pathogen. For the CLK-seq assay, we used rapid automated dosing system for rapid and efficient mixing of the crosslinking agent for a variable amount of time.



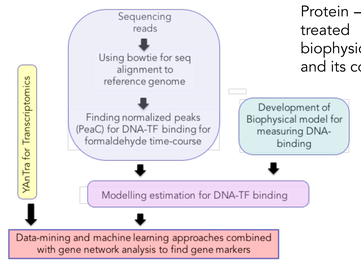
- Without the micro-mixer, water and bead solution remains separated at low (laminar separation) and high flow rates (hydrodynamic focusing).
- With the micro-mixer, the chaotic flow field enables efficient cross flow for rapid mixing.



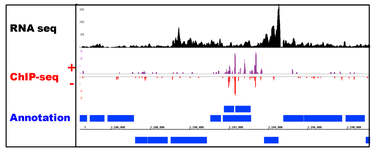
- The assembled dosing apparatus consists of 1 cross-linking station for rapid cross-linking (sub-second time scale) and 4 quenching stations for quenching the cross-linked mixture at different time points.
- Different quenching time points can be obtained by simply changing the tubing length.
- Ability to obtain multiple time data sets from a single experiment.



Strategy and Tools

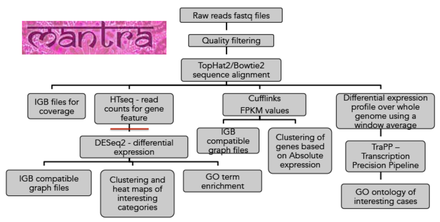


Protein – DNA interaction captured in *Klebsiella pneumoniae* treated with sublethal dose of antibiotics. We apply biophysical models to explain transcription factor binding and its correlation with transcription

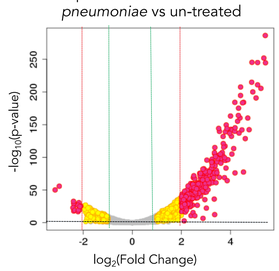


Research strategy used in the analysis of CLK-seq datasets. Different sections correspond to the different stages in the development of assay. At the top, grey section shows the initial processing of the sequencing data of the different crosslinking time points of DNase-seq assay. The yellow box for RNA-seq analysis. The blue section is the development of the kinetic model. Pink section denoted the development of software for model fitting and parameter estimation and finally the salmon color box for integration and data mining of the collected datasets for interesting biology and interpretation

RNA-seq Analysis by MAnTra (Multi Analysis Transcriptomics pipeline)

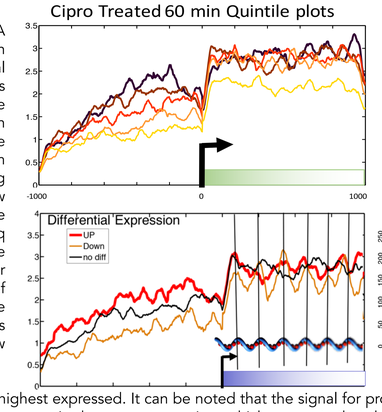


Ciprofloxacin Treated *K. pneumoniae* vs un-treated



Results: CLK-seq reveals that DNA activity and transcriptional response are correlated

Cipro Treated 60 min Quintile plots

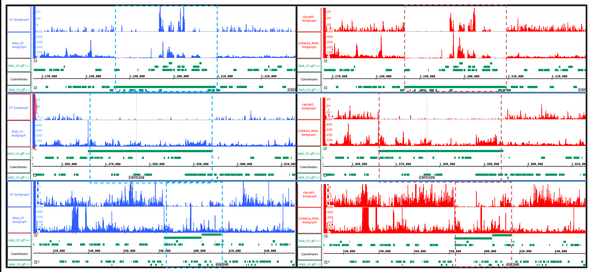


The protein-DNA interactions increase with the transcriptional response. For this analysis, genes in the genome were split in quintiles based on gene expression measured from RNA-seq analysis using MAnTra pipeline (from low to high). The average profile of CLK/DNase-seq read counts for these genes at the promoter and within the body of genes is shown in the figure. These five quintiles are shown from yellow being the genes least expressed to black being highest expressed. It can be noted that the signal for protein-DNA increases as the transcriptional response increases in the promoter region, which suggests that the regulatory activity of the genes with high transcriptional response is higher. In the gene body, only the low expressed genes (yellow) show a depletion in signal than the others groups. Also, the profile for DNA-protein interaction shows a dip in the signal at the gene start site. This profile resembles the Nucleosome-free region in Eukaryotes.

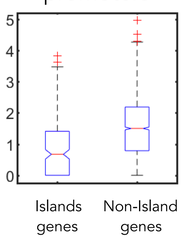
Average profile for DNase-seq for differentially expressed genes in presence of drugs is shown in the bottom figure. These average plots highlighted an interesting pattern in the differentially downregulated genes. These genes have a periodic accumulation of signal in the gene body at an amplitude of 100-150 nt. This pattern also resembles profile for nucleosome protection assay where the DNA locations associated with nucleosome are protected in Eukaryotic organisms. We speculate that these patterns are related to DNA compaction proteins which are associated with gene repression. Further experiments and studies are needed to explore this possibility

Results: Genomic Islands have fewer DNA activity recorded

CLK-seq and RNA-seq read maps at genomic islands



log_2(counts) at promoters



Using Genome Browser we noticed that for CLK-seq/DNase-seq read counts for genomic islands were depleted. In the six figures showed in top figure, the top track is DNase/CLK-seq data, the middle is transcriptomics data and bottom track shows genes. The boxed region shows the genomic islands. In all the twelve genomic islands the signal for CLK-seq is depleted. These regions show a great abundance of promoters and transcription but significantly low CLK-seq data. This indicates that either there is lack of DNA compaction in these regions or not the same level of regulation as non-mobile elements.

Discussion and Acknowledgment

We have developed a novel assay CLK-Seq which has proved to be a valuable tool in studying prokaryotic gene regulation. Preliminary data for CLK seq have revealed new observations in transcriptional regulation mechanisms for prokaryotic genomes and discovered new evidence of bacterial chromatin and its effect on gene regulation which was previously never been studied at the global scale. Overall applying CLK-seq and RNA-seq to study a MDR pathogen in great detail has shed more light on the pathogens ability to quickly respond to the drugs. Further work is needed to strengthen our claims and solidify our findings of early regulators as targets for the development of strategies to counter drug resistance.

This work was supported by Sandia's LDRD program.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

kpoorey@sandia.gov