Jeremy Hogaboom

July 20, 2020

**A Study of E. Coli RNA polymerase - Template interaction utilizing**

**CRISPR/dCas9 protein**

 The goal of FURSCA this summer was to model the interaction between *Escherichia coli* RNA polymerase (Eco RNAP), a DNA template, dCas9 protein, and guide RNA (sgRNA). This project was based on the idea of developing a reversible and specific transcriptional control mechanism via the protein dCas9 (Widom et al. 2019). To begin, I had to learn about molecular modeling, and familiarize myself with a wide range of software programs commonly used in the field of biochemistry. Then I would use these programs to model the interaction and make predictions about how and where the proteins would collide. This information would prove very useful when we could continue lab work. A key step of the lab work is to crosslink the proteins and knowing potential crosslinking locations would give us a leg up come fall. I am quite new to this project and the Rohlman lab. My personal goal this summer was to better familiarize myself with the Eco RNAP, and read current, related literature.

 As the summer comes to a close, I was able to get much of this done; although there is still a lot more to do. There were a lot of great structural data already available for our proteins. Eco RNAP has a lot of literature and solved structures to explore. I found several that I liked using the protein database RCSB (rcsb.org). By exploring the protein database, I also stumbled upon some amazing resources about the structure and function of Eco RNAP, and dCas9 (Kang et al. 2004, Chen et al. 2019). While the individual structures were available, the goal was to model them together. To do this, I utilized modeling programs such as UCSF Chimera, and PyMOL (Pettersen et al. 2004). Both programs were frequently mentioned in the literature that we read. I learned a lot about the tools and features that both programs had to offer. Chimera was easy to use and made looking at specific chains simpler. PyMOL had better tools for sequence mutation, and alignment that allowed me to combine different PDB’s and make pretty pictures for visualization. Using PyMOL’s alignment tool, I was able to narrow down potential crosslinking locations (figure 1). Unfortunately, I was not able to narrow them down enough to use any of the docking functions natively available in these programs, or dedicated docking servers such as ClusPro (Kozakov et al. 2017). I did make steady progress, and I am confident that with more time I will be able to narrow down potential locations enough to make use of these tools.

 My personal goal this summer was extremely successful. I have a wonderful reference list that I can use both during research and for my thesis. I learned a lot about the structure of Eco RNAP, and the function of the different subunits (Sutherland and Murakami, 2018). There is also a lot of work exploring protein-protein interactions that translates very well to our work (Mosrin-Huaman et al. 2004). I also was able to read about other work pausing transcription (Strobel et al. 2020). This allowed me to reexamine the paper that started our project with fresh eyes (Widom et al. 2019). Towards the end of the summer, we were even exploring some of the long-term benefits of reversible transcription control. The ability to tag and explore the spliceosome more effectively is one example. After this summer, I am more comfortable in my knowledge of the project and why we are doing it. I think that this better understanding will help me to design better experiments in the lab, and better utilize the data that we get from the benchwork.

 Overall, this summer was a fantastic, if somewhat abnormal introduction to research. I was able to learn a lot about the research topic, and further the work from home. Outside of the research, I personally learned a lot about staying focused and driven in a time and place that made that difficult. The ability to be flexible and productive in spite of strange circumstances is certainly invaluable. I struggled a lot with this when I was at high school and being able to overcome that at home was a big personal challenge for me. I am very thankful to FURSCA and my mentor Dr. Rohlman for allowing me to continue this work both personally and professionally from home. I am not sure what I would have done had the program been cancelled entirely. Covid FURSCA was a great experience, and I am excited to see what a more normal version of FURSCA will look like next summer.

Figure 1.

This figure shows the E. Coli RNA Polymerase (In green) aligned along the U5 DNA Template (Orange) approaching a bound dCas9 blockade (Blue). The red highlights are what I am currently looking at as potential crosslinking locations. They are still too large, but hopefully can be narrowed down further with future work.

References

(1)

Sutherland, C.; Murakami, K. S. An Introduction to the Structure and Function of the Catalytic Core Enzyme of Escherichia Coli RNA Polymerase. *EcoSal Plus* **2018**, *8* (1). <https://doi.org/10.1128/ecosalplus.ESP-0004-2018>.

(2)

Strobel, E. J.; Lis, J. T.; Lucks, J. B. Chemical Roadblocking of DNA Transcription for Nascent RNA Display. *J. Biol. Chem.* **2020**, *295* (19), 6401–6412. <https://doi.org/10.1074/jbc.RA120.012641>.

(3)

Chen, W.; Zhang, H.; Zhang, Y.; Wang, Y.; Gan, J.; Ji, Q. Molecular Basis for the PAM Expansion and Fidelity Enhancement of an Evolved Cas9 Nuclease. *PLOS Biology* **2019**, *17* (10), e3000496. <https://doi.org/10.1371/journal.pbio.3000496>.

(4)

Kang, J. Y.; Olinares, P. D. B.; Chen, J.; Campbell, E. A.; Mustaev, A.; Chait, B. T.; Gottesman, M. E.; Darst, S. A. Structural Basis of Transcription Arrest by Coliphage HK022 Nun in an Escherichia Coli RNA Polymerase Elongation Complex. *eLife* **2017**, *6*, e25478. <https://doi.org/10.7554/eLife.25478>.

(5)

Kozakov, D.; Hall, D. R.; Xia, B.; Porter, K. A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S. The ClusPro Web Server for Protein–Protein Docking. *Nature Protocols* **2017**, *12* (2), 255–278. <https://doi.org/10.1038/nprot.2016.169>.

(6)

Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res* **2000**, *28* (1), 235–242. <https://doi.org/10.1093/nar/28.1.235>.

(7)

Mosrin‐Huaman, C.; Turnbough, C. L.; Rahmouni, A. R. Translocation of Escherichia Coli RNA Polymerase against a Protein Roadblock in Vivo Highlights a Passive Sliding Mechanism for Transcript Elongation. *Molecular Microbiology* **2004**, *51* (5), 1471–1481. <https://doi.org/10.1111/j.1365-2958.2003.03926.x>.

(8)

Pettersen, E.; Goddard, T.; Huang, C.; Couch, G.; Greenblatt, D.; Meng, E.; Ferrin, T. UCSF Chimera--a Visualization System for Exploratory Research and Analysis. *J Comput Chem* **2004**, *25* (13), 1605–1612. <https://doi.org/10.1002/jcc.20084>.

(9)

Widom, J. R.; Rai, V.; Rohlman, C. E.; Walter, N. G. Versatile Transcription Control Based on Reversible DCas9 Binding. *RNA* **2019**, *25* (11), 1457–1469. <https://doi.org/10.1261/rna.071613.119>.