CONSIDER FOR TALK

2025 SEA Symposium Abstract

University of South Florida

Tampa FL

Corresponding Faculty Member: Richard Pollenz (pollenz@usf.edu)

Use of AlphaFold to Predict Dimerization and DNA-Binding of Four Novel HTH DNA Binding Proteins from the Immunity Cassette of Arthrobacter Phage Alatato

Alvin Varghese, Shilen Abraham, Nihal Anantoji, John Coleman, Rebecca Aguilera, Marco Barillas, Ryan Barstys, Emily Benson, Kayce Boucher, Adrianna Chacko, Niveditha Chandrakanth, Angel Chau, Angela Chen, Luong Tri Duc, Olivia Godfrey, Sasha Hoang, Medhana Kethamreddy, Nisarg Kumar, Bao Ngan Le, Gregory Lin, Kim Nguyen, Madhura Pandit, Sarah Robinson, Kelie Shah, Mohini Sharma, Sophia Singh, Josh Smith, Erin Stanford, Lucy Thomas, Ian Wells, Wynter Dean

Alatato infects Arthrobacter globiformis and grows with a temperate life cycle as evidenced by a cloudy plaque phenotype. Temperate phages contain genes for an integrase, immunity repressor, cro, and excise and in many phages, these genes are found in a defined immunity cassette with the integrase followed by a reverse-forward gene transition with the immunity repressor (rev) and cro (for). The excise is generally downstream of the cro. The immunity repressor and cro are DNA binding regulatory proteins with an N-terminal HTH motif for DNA binding and a C-terminal domain for dimerization. Alatato contains a tyrosine integrase (gene 29) and genes 31 (rev), 32 (for) 33 (rev) and 34 (for) all encode proteins of <95aa that have defined HTH domains. gp31, gp33 and gp34 hit to lambda CI PDB 7JVT\_D with >97% probability. Forward gene 35 encodes the excise, thus, genes 29-35 define the Alatato immunity cassette. However, since genes 31-34 all have defined HTH domains and generally hit to the same targets, HHpred could not be used to assign repressor or cro function. AlphaFold and ChimeraX were used to model the four HTH proteins and evaluate their ability to 1) form homodimers, 2) form heterodimers and 3) bind to predicted regulatory DNA sequences. All HTH proteins could be folded as monomers at high confidence, but none were predicted to bind DNA. AlphaFold-Multimer was used to assess dimerization and binding to 50bp DNA sequences. AlphaFold assigns an interface-predicted template modelling score that measures the accuracy of the predicted interface between the subunits of the protein-protein or protein-DNA complex. Scores of >0.8 show that the predicted interaction is highly confident. gp31:gp31 (0.79), gp33:gp33 (0.87), gp34:gp34 (0.86), gp31:gp33 (0.8), gp32:gp33 (0.93), and gp32:gp34 (0.89) all showed high confidence predictions. The interactions of all dimers by both AlphaFold and ChimeraX revealed that the predicted dimerization domain was in the C-terminal region of the monomers and that the HTH region was exposed for possible DNA binding. When the various dimers were modeled with 50bp of upstream DNA to genes 30, 31, 33 and 34, the data showed that the HTH region was predicted to contact the DNA. The highest confidence was found for gp31:gp33 binding to gene 31 sequence (0.80) and gp32:gp33 binding to gene 33 sequence (0.82). The gp33:gp33 bound to upstream gene 34 DNA at 0.73 and the gp34:gp34 bound to upstream gene 33 DNA at 0.66. Collectively, these results confirm that these novel HTH proteins are predicted to function as dimers even though they are <100 amino acids and lack the canonical dimerization domain of lambda C1 repressor. The DNA binding data was not as robust but suggests that the gp33 homodimer may bind to DNA upstream of gene 34 and the gp34 homodimer may bind to DNA upstream of gene 33. This data can now be used to design wet lab experiments to assess how these dimers function to regulate the genetic switch in FB cluster phages.