Iowa Initiative for Artificial Intelligence
Final Report

Project title:	Predicting interactions between extracellular lipases and endogenous inhibitors
Principal Investigator:	Brandon Davies and Michael Schnieders
Prepared by (IIAI):	
Other investigators:	
Date:	6/30/22

Were specific aims fulfilled:	Aim 1: Yes Aim 2: partially
Readiness for extramural proposal?	N
If yes … Planned submission date	
Funding agency	
Grant mechanism	
If no … Why not? What went wrong?	Further progress on aim 2 is needed to identify interacting residues

Brief summary of accomplished results:

Misregulation of lipid metabolism is associated with a variety of metabolic diseases, including coronary heart disease and atherosclerosis. Three members of the angiopoietin-like family of proteins, ANGPTL3, ANGPTL4, and ANGPTL8, are known to regulate both plasma triglyceride and cholesterol levels by inhibiting two important lipases, lipoprotein lipase, which mediates clearance of plasma triglycerides, and endothelial lipase, a regulator of HDL. Despite the importance of all these proteins in regulating lipid metabolism, only the structure of lipoprotein lipase has been solved. The structures of the endothelial lipase and the ANGPTL proteins are not know, nor are their interactions with each other understood. In this project we aimed to model the structures of these proteins and predict key molecular interactions between them. We were successful in modeling the structures of ANGPTL3, ANGPTL4, ANGPTL8, and endothelial lipase monomers, as well as the more physiologically relevant ANGPTL3 trimers and ANGPTL3-ANGPTL8 complexes. Combining these structures with wet-lab mutagenesis studies suggested that certain mutants disrupted ANGPTL3 function by disrupting ANGPTL3 trimer formation. Using wet-lab experiments we were able to confirm that this was indeed the case.

Research report:

Aims (provided by PI):

Specific Aim 1: Generate predicted structures of ANGPTL3, ANGPTL4, ANGPTL8, and endothelial lipase.

Specific Aim 2: To predict the amino residues critical for the interactions between ANGPTL proteins and their target lipases.

Data:

The input data for structure predictions were the amino acid sequences of the proteins of interest (Endothelial lipase, ANGPTL3, ANGPTL4, ANGPTL8). For multimeric chains the inputs included multiple protein sequences or multiple copies of the same protein sequence. For mutational studies, the

inputs were .pdb files of the predicted structure of the ANGPTL3 trimer or the ANGPTL3-ANGPTL8 complex.

AI/ML Approach, Experimental methods, validation approach:

To generate structures, we initially used trRosetta to model structures using amino acid sequence. Shortly after beginning these studies, Alphafold2, an AI approach to predicting protein structure developed by Deepmind, as well as the structures generated by this approach went public (1, 2). We then used these structures as a basepoint for further analysis. We also used two Alphafold variants, Colabfold (3) and Alphafold multimer (4) to generate predicted structures of ANGPTL3 trimers, and ANGPTL3-ANGPTL8 complexes. In conjunction with these computationally generated structures, we performed wet lab experiments, mutating residues of ANGPTL3 and examining the functional consequences. Computational analysis of point mutants was performed using Dynamut2 (5), a machine learning tool for estimating the change in structural free energy incurred by point mutations.

Results:

Specific Aim 1: Generate predicted structures of ANGPTL3, ANGPTL4, ANGPTL8, and endothelial lipase.

Initially, we generated structures using TrRosetta (6)(data not shown). However, shortly after beginning these studies, Alphafold2 from Deepmind and the structures it generated went public (1, 2). We retrieved the structures of endothelial lipase, ANGPTL3, ANGPTL4, and ANGPTL8 from the Alphafold2 database (Fig. 1). Endothelial lipase is likely active as a monomer, and as expected the predicted structure was similar to that of lipoprotein lipase, a closely related lipase that is also a target of ANGPTL proteins (7, 8). However, the physiological functional structures for ANGPTLs are oligomeric. ANGPTL3 is know to oligomerize and is thought to form trimers (9, 10). ANGPTL8 must form a complex with ANGPTL3 to be functional (11–14), and ANGPTL4 is though to form a tetramer (15, 16). As Alphafold is not able to generate multimeric structures, we turned to two additional machine-learning tools to generate these complexes, Colabfold (3) and Alphafold multimer (4). Both methods are limited in that the size of complexes modeled is limited by the available computer memory. Because of this limitation, for complexes, we modeled only the N-terminal domain of ANGPTL3 and ANGPTL4, as this domain is sufficient and necessary for inhibition of lipase targets. Using this approach, we generated predicted complex

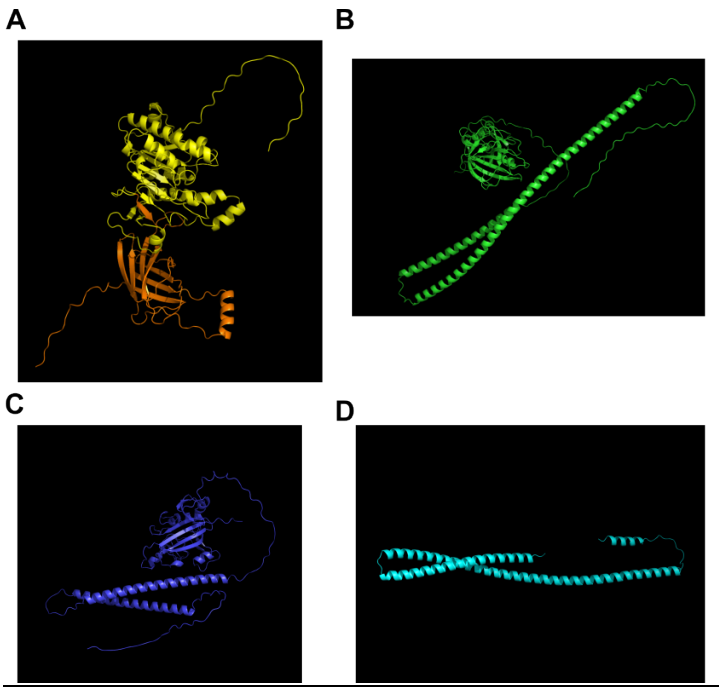

Fig 1. Alphafold predicted structures of EL, ANGPTL3, ANGPTL4, and ANGPTL8. Alphafold generated structures of human endothelial lipase (A), ANGPTL3 (B), ANGPTL4 (C), and ANGPTL8 (D).

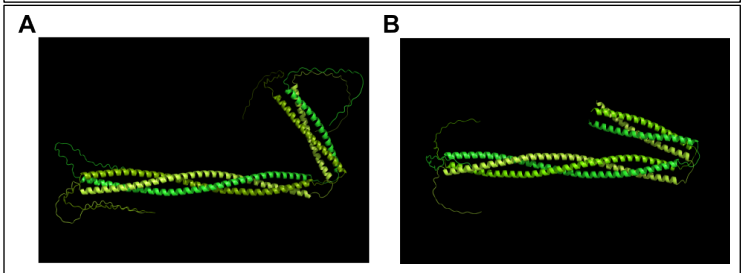

Fig 2. Predicted structures of ANGPTL3 trimers. CollabFold generated structures of human (A) and mouse (B) ANGPTL3 trimers. Different shades represent the 3 ANGPTL3 chains.

structures of human and mouse ANGPTL3 trimers (Fig. 2) and ANGPTL4 tetramers (data not shown). The N-terminal domain of ANGPTL3 has been predicted to form a coiled coil and our computationally generated models support that prediction (Fig. 2). Interestingly, when we attempted to generate models of the ANGPTL3-ANGPTL8 complex, the top predicted models separated into two main forms. In one form, ANGPTL8 formed a coil with the three ANGPTL3 chains throughout the N-terminal domain (Fig. 3A). In the second form, ANGPTL8 formed a coil with the ANGPTL3 chains for the more distal alpha helix, but did not interact with the larger alpha helix (Fig. 3B). Both forms were predicted for both mouse (Fig. 3) and human (data not shown) ANGPTL3-ANGPTL8 complexes.

Concurrent with the computation modeling of structures, our lab also started a mutagenesis study of ANGPTL3, in which residues of ANGPTL3 were mutated to alanine and the resulting proteins were tested for ability to inhibit lipoprotein lipase and endothelial lipase. When we mapped mutations onto the predicted structure of the ANGPTL3 trimer, we noticed that many of the mutations that disrupted the ability of ANGPTL3 to inhibit lipases were mapped to the interface the chains forming the coiled coil (Fig. 4). This observation led us to hypothesize that these mutations might not disrupt the interactions of ANGPTL3 and lipases per se, but rather might disrupt the proper formation of ANGPTL3 trimers. To test this hypostasis we ran wild-type and mutant proteins on blue native gels to assess the oligomeric state of ANGPTL3. In support of our hypothesis, we found that many of the mutations we tested, no longer formed trimers and were present primarily as monomers (Fig. 5)

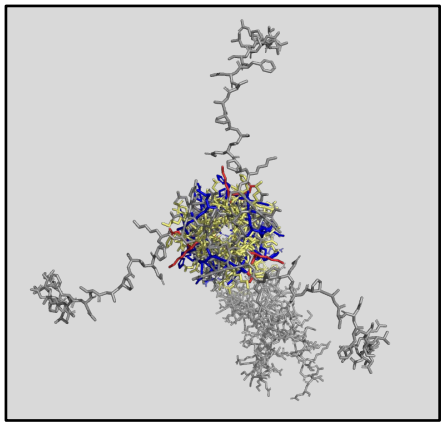

Fig 4. Mutations predicted to be at the coiled coil interface of ANGPTL3 disrupt ANGPTL3 function. Structure of mouse ANGPTL3 (shown in Fig. 2B) is oriented so that the interface of the coiled coil can be seen. Residues that when mutated disrupt the ability of ANGPTL3 to inhibit LPL and EL are shown in yellow.

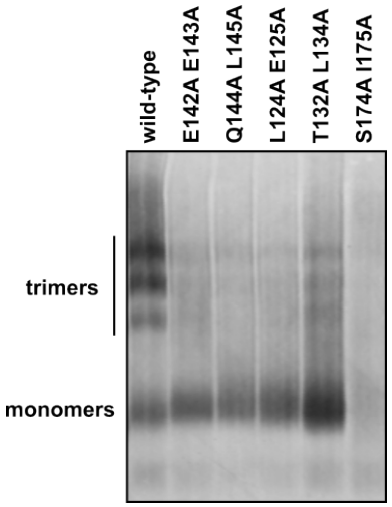

Fig 5. Effect of selected ANGPTL3 mutations on ANGPTL3 trimerization. Blue native gel showing the oligomerization state of wt ANGPTL3 and 5 ANGPTL3 mutants with impaired EL inhibition.

Specific Aim 2: To predict the amino residues critical for the interactions between ANGPTL proteins and their target lipases.

Our ultimate goal is to model the interactions between ANGPTL3 trimers and endothelial lipase and between ANGPTL3-ANGPTL8 complexes and lipoprotein lipase. The size of the relevant proteins and complexes makes it difficult to model these interactions with Colabfold or Alphafold-multimer. To avoid the computational limits of these methods we attempted to model the interactions of a trimer of ANGPTL3 N-terminal domains and the N-terminal domain of endothelial lipase (this is the catalytic domain of the protein). Although we were able to produces some predicted structures (Fig. 6), close examination of these models indicate a number of impossible steric clashes, undermining any

confidence in the output. The field of multimeric protein complex prediction is evolving rapidly and we hope that further optimization will allow us to produce better predictions of the interactions between ANGPTL proteins and their lipase targets.

As another approach to assess the residues of ANGPTL3 that are important to its function, we used DynaMut2, a machine-learning tool for assessed the impact of mutations on the free energy of protein structures (5). Using this tool, we have generated prediction for the free energy changes that would be caused by each of the mutations we plan to make in our mutagenesis study. As we collect more wet-lab data we will compare these predictions with actual changes in ANGPTL3 function and trimerization. If we can generate more confident predictions of endothelial lipase-ANGPTL3 interactions, this tool might also be useful in identifying residues most important for these interactions.

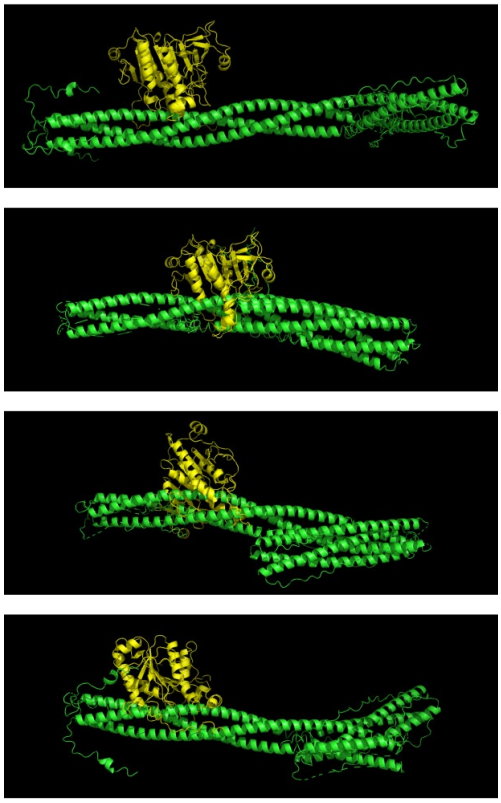

Fig 6. Four predicted structures of ANGPTL3 N-terminal trimers bound to the N-terminal domain of EL as predicted by Alphafold-Monomer. EL is shown in yellow and the ANGPTL3 trimer is shown in green.

Ideas/aims for future extramural project:

Combining computation/machine learning approaches with wet-lab mutagenesis studies has proven very informative. In future extramural proposals, we would like to expand this concept, using machine learning approaches to predict important residues on lipoprotein regulators, testing those predictions using mutagenesis and biochemical approaches, and then using the experimental results to further inform computational approaches. We recently received fundable scoring for a grant that includes the mutagenesis and biochemical arm of some of these studies and will incorporate the computational approaches into future proposals.

Publications resulting from project:

No publications have arisen from this project at this time, but we plan to submit a small study that includes a small part of this data in the next couple of months. A larger manuscript that includes the bulk of the data generated from this project as well as our mutagenesis study is planned for the next year.