Homology Modeling Studies to Structurally Characterize BaiK, a Proposed Coenzyme A Transferase

Roman Aguirre, Carolina Mata, Stephen Khuu, Kathryn McCulloch
Department of Chemistry and Biochemistry
Kellogg Honors College Capstone Project

Abstract

Primary bile acids, such as cholic acid, play a major role in digestion through the break-down and absorption of fats and oils that enter the body. Once in the gastrointestinal tract, naturally occurring gut bacteria such as *Clostridium hylemonae* convert primary bile acids into secondary bile acids and require a putative Coenzyme A transferase encoded within the bile acid inducible operon. For *in vitro* lab work, a recombinant approach using *E. coli* cell lines was taken. First, the *E. coli* Tuner expression line was transformed with a vector plasmid containing the *baiK* gene for overexpression. Large scale overexpression was used to harvest BaiK in large quantities, and the resulting combined cell pellet was collected for purification using an AKTA UltraPure instrument via affinity column chromatography. Resulting purified protein was analyzed via SDS-PAGE. Challenges in BaiK purification are being troubleshooted, and the protocol is still being optimized. Due to the ongoing pandemic, a transition into *in-silico* modeling was made. Homology modeling is being utilized to characterize and identify potential BaiK folding patterns and key active site residues using web-based tools. Models produced are being validated by using the built-in functions and information from modelling servers.

Background

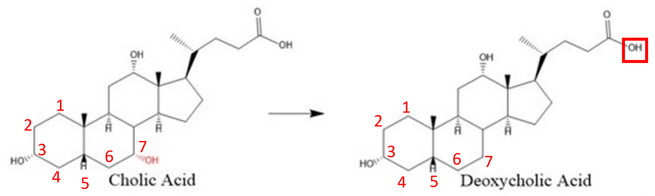

Figure 1: 7α-carbon dehydroxylation reaction that occurs when naturally occurring gastrointestinal fauna digest this hydroxyl group, resulting in secondary bile acids. The red box highlighting the hydroxyl group of the carboxyl is where the CoA molecule is attached, forming a thioester bond.

Primary bile acids play an important role in fat and nutrient absorption during metabolism. These primary forms, such as cholic acid **(Figure 1**) and chenodeoxycholic acid, are converted via a 7-alpha dihydroxylation to yield secondary bile acids. These secondary forms then interact with CoA transferases such as BaiK. This protein operates by facilitating the transfer of Coenzyme A, an important carrier molecule, onto a converted bile acid.

Methods- Homology Modeling

Preparation
1. ExPASy Translate
2. ExPASy Peptide Cutter

Sequence Alignment
1. NCBI BLAST tool
2. Cobalt Alignments
3. Espript 3 secondary structure alignment

Model and Validation
1. PYMOL and WinCoot
2. Swissmodel
   a. Built-in validation

Results

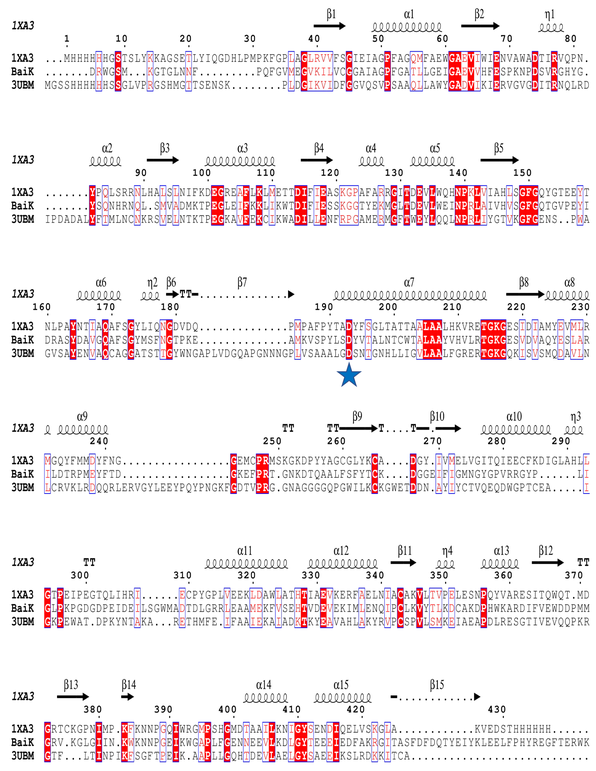

Figure 2: Sequence alignment between BaiK, and two other proteins obtained using the NCBI BLAST tool. These are described in **table 1**, shown below. The blue star denotes the key catalytic aspartic acid residue.

Table 1: Template proteins with their protein name and corresponding PDB code. Additionally, validation parameters and values for each homology model made using these proteins are provided

Protein PDB Code	Protein Name	GMQE	QMEAN	RMSD Score	Percent ID/similarity/gap
1XA3, model 1	CaiB	0.67	-1.60	0.070	36.8/5/3
3UBM, model 2	Formyl CoA: oxalate CoA Transferase	0.56	-2.37	0.394	28/44/13

Model 1

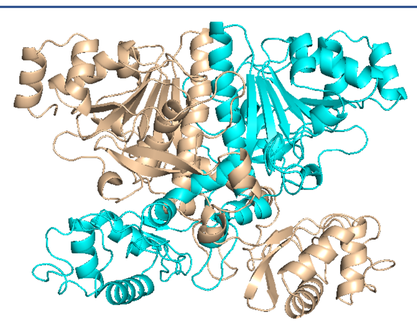

Figure 3: Homology model 1, which uses PDB protein 1XA3 as a template. The dimer displays an interlaced ring-like structure, which is theorized to be a trademark of type III CoA transferases.

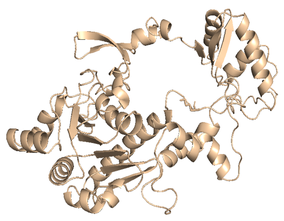

Figure 4: Monomer chain A of generated homology model 1

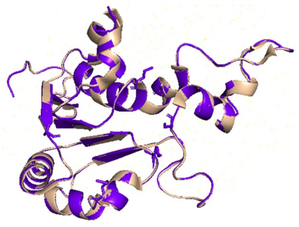

Figure 5: Overlapping between monomer chain A and its template protein 1XA3

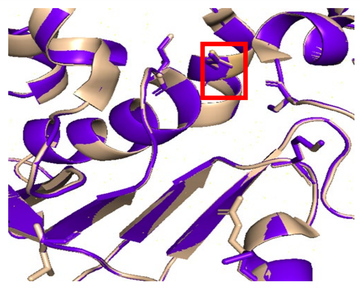

Figure 6: monomer active site, with key catalytic aspartic acid residue boxed in red. Aspartic acid plays a critical component in binding the Coenzyme A molecule in the large subunit.

Model 2

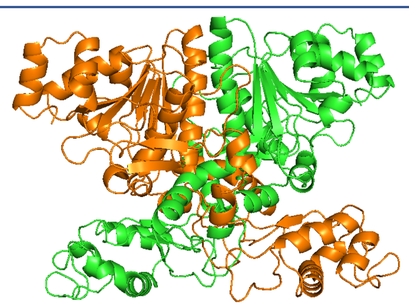

Figure 7: Homology model 2, which uses PDB protein 3UBM as a template. The dimer displays an interlaced ring-like structure, which is theorized to be a trademark of type III CoA transferases.

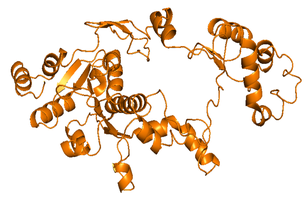

Figure 8: Monomer chain A of generated homology model 1

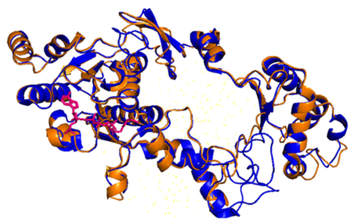

Figure 9: Overlapping between monomer chain A and its template protein 3UBM

Figure 10: monomer active site, with key catalytic aspartic acid residue boxed in red. Aspartic acid plays a critical component in binding the Coenzyme A molecule in the large subunit.

Dimer

Monomer

Active Site

Discussion and Validation

Homology models were evaluated using multiple criteria using built-in functions of the SWISS-model program. GMQE scores are numbered on a scale from 0 to 1, with scores <0.7 being considered to be in a favorable range. Homology models 1 and 2's score was within this range, with values of 0.67 and 0.56 respectively. Additionally, QMEAN scores for reliable models are expected to be within the range of 0 to -4. Both reported scores for Models 1 and 2 were also within this range, with values of -1.60 and -2.37 in that respective order. The placement in the middle of this given range may be an indicator that these generated models are semi-reliable. Another validation parameter used was the RMSD value, which is not a built-in function of SWISS-Model. This is a validation parameter of PyMOL visualizing software, in which the original template protein is aligned with the generated homology model, indicating the mean deviation of the model from its original template protein. Overall, the sequence alignment and validation methods show that there are sequence areas that are not well conserved between one model and its template protein. The large domain contains the most conserved sequence area, while the small domain is the main contributor to the uncertainty of these values. Additionally, a key catalytic residue was identified in the 176D position. For future studies, more catalytically-relevant residues should be identified for BaiK and should be compared to its template counterparts to better understand the structural and mechanistic components of the CoA transferase reaction.

Previous literature has hypothesized that type III CoA transferases are typically only dimers and are usually present in a "ring structure" in which one monomer is overlapped by the other. This can be observed with both homology models generated, visualized in **figures 3 and 7**. Furthermore, there are two main domains, a large and a small one, as seen in **Figure 11**. Consistent with other CoA transferases, both homology models generated show a Rossmann fold (**Figures 12**) in the large domain. The largest conserved sequence areas are those in the large domain. Therefore, it can be hypothesized that the large domain is the one that binds a CoA molecule, which explains the sequence consistency with other similar proteins. The small domain could be what binds the substrate specific to the enzyme, which for BaiK is a bile acid. In future studies, knockouts should be done to confirm the importance of key catalytic residues, such as 176D and their impact on enzymatic activity. Additionally, the monomeric unit should be further analyzed in order to understand the relevance of the ring structure in type III CoA transferases.

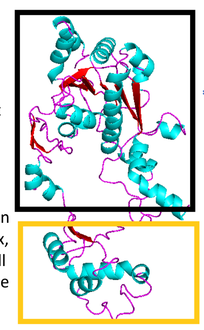

Figure 11: Chain A of homology model 1, which is split into its large domain, containing the CoA binding site in the black box, and the small domain in the yellow box

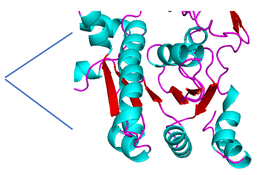

Figure 12: Close-up shot of the large domain Rossmann fold. This type of fold is a conserved motif in type III CoA transferases.

Acknowledgements

Dr. McCulloch has received past support from CSUPERB via a New Investigator award and has currently received support from the Cal Poly RSCA program. A special thank you to Dr. McCulloch for helping us shift the direction of the project in a way that allowed continuation of research during the COVID-19 pandemic

References

Rangarajan, E.s., et al. "Crystal Structure- P1 Form- of Escherichia Coli Crotonobetainyl-CoA: Carnitine CoA Transferase (CaiB)." *American Chemical Society*, 2005, doi:10.2210/pdb1xk6/pdb.

"CoA-Transferases." *CoA-Transferases - an Overview | ScienceDirect Topics*, www.sciencedirect.com/topics/agricultural-and-biological-sciences/coa-transferases (accessed March 4, 2020)

Lee, Sangbae, et al. "Sampling Long Time Scale Protein Motions: OSRW Simulation of Active Site Loop Conformational Free Energies in Formyl-CoA:Oxalate CoA Transferase." *Journal of the American Chemical Society*, vol. 132, no. 21, Feb. 2010, pp. 7252–7253., doi:10.1021/ja101446u.

Vital, Marius, et al. "Diversity of Bacteria Exhibiting Bile Acid-Inducible 7α-Dehydroxylation Genes in the Human Gut." *Computational and Structural Biotechnology Journal*, vol. 17, 2019, pp. 1016–1019., doi:10.1016/j.csbj.2019.07.012.