Group Members: Anusheh Chaudry & Aakriti Lakshmanan

Introduction:

We were interested in examining how genetic variants in CYP3A5 and ABCB1 affect the efficacy of the Tacrolimus drug. Tacrolimus is an immunosuppressant that prevents organ rejection after transplantation. This drug has a narrow therapeutic range, which means that the range between when it is effective and when it is toxic is smaller when compared to other drugs (Haverals et al.). Furthermore, genetic variations can play a larger role in tacrolimus absorption and efficiency - what the "correct" dosage of the drug varies from person to person. We specifically were interested in examining known gene-drug interactions from the PGKB database, and simulating how mutations in these genes can impact the way tacrolimus binds, thereby impacting efficacy.

Background:

Past research on gene-drug interactions with impact on efficacy was found by examining the variant annotations in the PGKB database. There were two main genes represented in this dataset, CYP3A5 with 586 annotations, and ABCB1 with 101 annotations.

ABCB1 is a gene that codes for a protein in the ATP-binding cassette transporters family. These transporters move proteins across cell membranes, and the specific protein that this gene codes for, P-glycoprotein (P-gp), is involved in multidrug resistance. This protein is an energy-dependent efflux pump that actively transports a wide variety of substances and is expressed in tissues such as the intestines, liver, kidneys, and the blood-brain barrier. Within the PGKB database, there were 5 specific variants associated with a difference in tacrolimus efficacy based on past research: rs1045642, rs1128503, rs2032582, rs2229109, and rs9282564. We decided to look further into the rs2229109 and rs2032582 since these polymorphisms lead to a structural change in P-glycoprotein and would allow us to explore the physical differences that

these changes create. rs2229109 is a single nucleotide polymorphism (SNP) that leads to an amino acid substitution from Ser to Asn in the ABCB1 protein. Past research has found that this change impacts the transport of tacrolimus, which may in turn impact efficacy (Dessilly et al.). rs2032582 is a polymorphism that is also known as G2677T/A - at position 2677 in the ABCB1 gene, the nucleotide can be either guanine (G), thymine (T), or adenine (A). Individuals with this mutation showed higher concentrations-to-dose ratios, as well as an increase in the risk of sepsis (Brunet and Pastor-Anglada) (Azam et al.).

CYP3A5 is a gene that encodes a protein in the cytochrome P450 family, specifically involved in the metabolism of a wide range of drugs. It is part of the primary enzymatic system responsible for metabolizing nearly half of all prescribed medications. Past research has found CYP3A5 to have a major impact on the metabolism of tacrolimus. Due to this, the tacrolimus dosing equation has been developed to include the CYP3A5 genotype as a factor for determining dosage of the drug. Individuals with the CYP3A5*3 genotypes exhibit slower drug metabolism and thus require lower tacrolimus doses. On the other hand, individuals with the CYP3A5*1 genotype metabolizes tacrolimus more efficiently, and thus requires higher tacrolimus doses to achieve therapeutic levels. Due to this we decided to look further into CYP3A5*1 and CYP3A5*3 since they directly relate to dosage of tacrolimus.

Past research has focused primarily on clinical studies and examining how individuals with a specific genotype react to tacrolimus. This research has enabled others to utilize genotype-guided models to find the correct personalized dosage plan for tacrolimus (Shi et al.). However, we were unable to find structures for the variants mentioned above, and were interested in understanding the structural mechanisms behind these variants.

Methods:

We first determined the sequences of the variants by taking the wild-type sequence from UniProt and applying the mutations to the sequence. For example, for rs2229109, we swapped the amino acid at position 400 from Serine to Asparagine. This sequence was fed into the AlphaFold server, and the first structure prediction was downloaded and converted from . cif to .pdb. We loaded these mutated structures and the wild-type structure from PDB (6GDI for ABCB1 and 5VEU for CYP3A5) into PyMol, and aligned them to understand the structural differences that these mutations caused.

Next, we performed molecular docking using AutoDock Vina to dock the drug to both wild-type and mutant protein structures. We used flexible docking, and chose the parameters of the process by consulting relevant literature. For ABCB1, there is a large binding cavity of size 6000 cubic amstrongs, defined by the transmembrane helices (Gutmann et al.). We determined which parts of the protein to designate as a ligand docking area by downloading annotations from PDBTM, and filtering for residues that are designated as transmembrane alpha helix. For CYP3A5, the binding cavity was defined to be around the heme group and specifically targeted key active site residues that included residues like lle118, lle301, and Phe304, all of which are key for drug metabolism. The docking parameters for CYP3A5 were based on a recent study, however for CYP3A5*3 there were adjustments made with the flexible residue selection, shifting from A:304 to A:120, to allow the residue to be closer to the truncated region and more representative of structural changes (Islam et al.).

Results and Analysis:

AlphaFold uses deep learning to predict protein structures, and has the highest accuracy currently when compared to other deep learning methods such as ColabFold or COSMIC2. It also uses patterns from known protein data, which makes it useful for modeling genetic

variations. AutoDock Vina simulates the ligand binding/docking by examining the most favorable binding conformations and estimated energies. It uses the vina scoring function to predict binding affinity. Using these tools together allows us to further examine the question of how genetic variations may affect tacrolimus binding and transport.

There are other approaches that could have worked, including using ColabFold or COSMIC2 for predicting the structure of the variants or using GLIDE or GROMACS to simulate ligand docking. We found that AlphaFold, especially using the server, was really easy to use and quicker with no local installation needed. Similarly, the ease of use of the Colab guides for Autodock Vina enabled us to quickly test multiple genes and multiple variants. We used these two methods combined because it provided an optimal balance between accuracy and efficiency.

For ABCB1, we saw that the binding affinity for rs2229109 was slightly better than the affinity for the wild-type ABCB1 (-6.744 vs -6.705). We found this result to be interesting because the predicted AlphaFold structure for rs2229109 was also the most different than the other two variations tested, with an RMSD of 8.296. The variant annotation in PKGB mentioned that individuals with this genotype had a higher chance of allograft loss (Woillard et al.). Since ABCB1 is a efflux pump, a higher binding affinity could mean that the drug is effluxed efficiently and quickly from the body, reducing the efficiency of tacrolimus itself (Dessilly et al.). However, past research has also found that the 1199A variant protein has nearly no effect on tacrolimus intracellular accumulation. The significant structural changes within ABCB1 may also lead to the allograft loss. For both the variations of rs2032582, we saw that the binding affinity was worse with values of -6.44 and -6.552 for the S to A and S to T variations respectively. We also saw that the predicted protein had a structure that varied more greatly from the wild-type protein, with extra solution-adjacent loops that aren't present in the wild-type

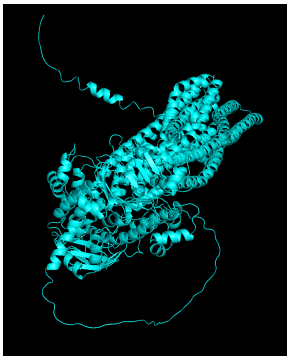
structure. The structures are also much less symmetric, and we were surprised to see that tacrolimus binds to very different regions of the ligand. This might be due to the stochastic nature of the docking software. Most of the annotation in PKGB looked at this mutation in conjunction with other mutations, including those on CYP3A5. A lower binding affinity and wrong location of binding might lead to overaccumulation of the drug within cells, causing toxicity and other side effects.

For CYP3A5,we found that the binding affinity for CYP3A5*1 was slightly better than the affinity for the wild-type CYP3A5 (-7.206 vs -7.018). We can notice that the predicted CYP3A5*1 structure is mostly similar to the wild-type structure, with only a slight difference in the presence of helices near the top of the structure. On the other hand, we found that the binding affinity for CYP3A5*3 was much lower than the affinity for the wild-type CYP3A5 (-4.934 vs -7.018). This was expected, as there was a large difference in the predicted structure from the wild type, since the CYP3A5*3 mutation introduces a cryptic splice site in the gene, leading to altered mRNA splicing, a frameshift, and the production of a truncated, nonfunctional protein. Both of these findings fall directly in line with what we might expect due to results from past findings. CYP3A5*3 has a lower binding affinity, which would indicate that the genotype would exhibit slower drug metabolism to avoid drug accumulation and toxicity. On the other hand, CYP3A5*1 metabolizes tacrolimus more efficiently, resulting in faster drug clearance and thus requires higher tacrolimus doses to achieve therapeutic levels, as their metabolism reduces drug concentrations more quickly.

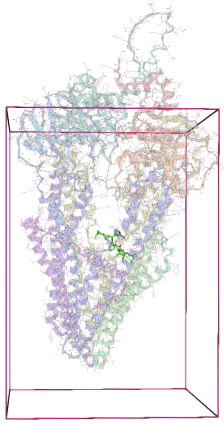
ABCB1:

rs2229109 (S -> N)	rs2032582 (S -> A)	rs2032582 (S -> T)	wild-type

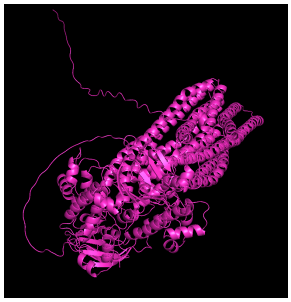
Group Members: Anusheh Chaudry & Aakriti Lakshmanan



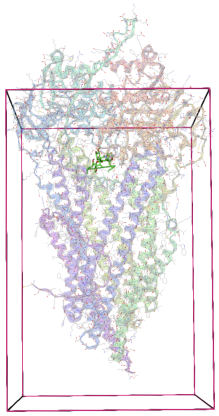
Rmsd = 8.296



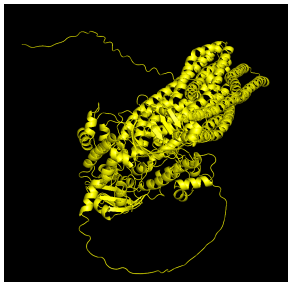
Mode	Affinity (kcal/mol)	RMSD l.b.	RMSD u.b.
1	-6.744	0	0
2	-5.337	1.92	3.352
3	-5.312	4.722	11.73
4	-5.22	4.752	12.08
5	-5.113	35.45	49.08
6	-5.002	50.03	66.11
7	-4.961	24.17	28.25
8	-4.917	23.5	30.3
9	-4.914	35.51	48.97



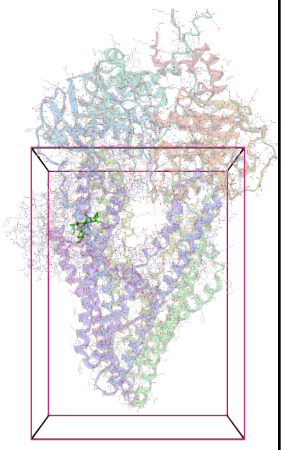
Rmsd = 6.742



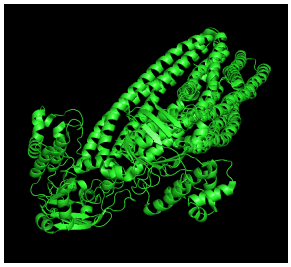
Mode	Affinity (kcal/mol)	RMSD l.b.	RMSD u.b.
1	-6.44	0	0
2	-6.408	1.433	2.208
3	-6.071	1.803	2.53
4	-5.843	19.99	25.23
5	-5.631	29.54	33.8
6	-5.591	4.525	11.16
7	-5.446	56.31	65.32
8	-5.419	14.74	19.7
9	-5.367	16.32	19.12



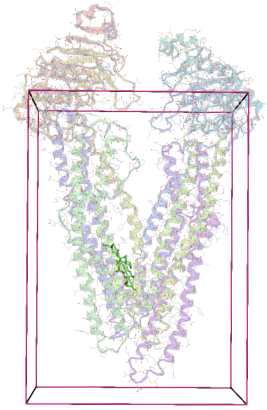
RMSD = 7.096



Mode	Affinity (kcal/mol)	RMSD l.b.	RMSD u.b.
1	-6.552	0	0
2	-5.959	1.297	2.026
3	-5.791	17.98	23.68
4	-5.629	17.98	22.78
5	-5.552	31.16	40.79
6	-5.375	2.397	5.344
7	-5.326	1.854	3.709
8	-5.291	25.05	31.91
9	-5.206	2.896	5.777

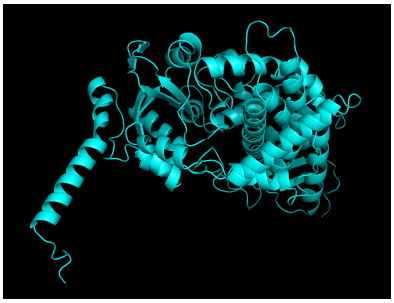
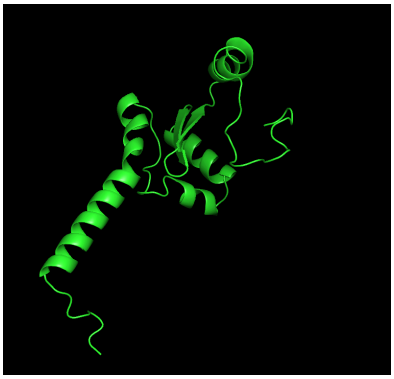
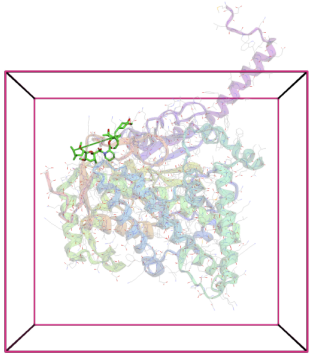
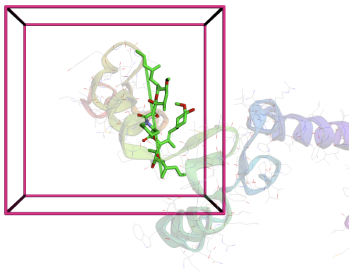
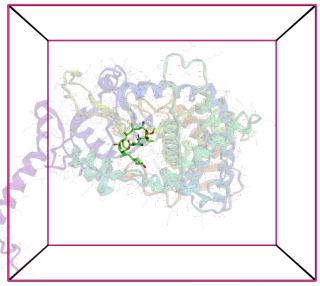


N/A



Mode	Affinity (kcal/mol)	RMSD l.b.	RMSD u.b.
1	-6.705	0	0
2	-6.541	47.7	56.13
3	-6.295	2.18	4.454
4	-6.26	46.52	54.84
5	-6.084	2.446	4.05
6	-5.927	46.65	54.98
7	-5.849	6.98	12.72
8	-5.772	6.406	11.93
9	-5.768	6.107	11.27

CYP3A5:

CYP3A5*1	CYP3A5*3	wild-type
		
Rmsd = 0.142	Rmsd = 0.631	N/A
		

Mode	Affinity (kcal/mol)	RMSD l.b.	RMSD u.b.
1	-7.206	0	0
2	-7.083	21.78	28.07
3	-7.056	1.41	2.057
4	-6.794	2.385	5.113
5	-6.724	2.754	11.98
6	-6.651	2.114	4.294
7	-6.506	2.244	5.163
8	-6.501	1.81	4.953

Mode	Affinity (kcal/mol)	RMSD l.b.	RMSD u.b.
1	-4.934	0	0
2	-4.931	4.229	8.186
3	-4.856	2.083	4.597
4	-4.821	5.189	8.44
5	-4.821	2.018	3.249
6	-4.81	4.063	5.932
7	-4.686	2.289	10.41
8	-4.674	4.439	8.168

Mode	Affinity (kcal/mol)	RMSD l.b.	RMSD u.b.
1	-7.018	0	0
2	-6.269	2.309	8.724
3	-5.948	16.56	22.87
4	-5.87	13.05	19.68
5	-5.861	18.07	24.99
6	-5.796	13.47	20.04
7	-5.647	13.46	19.25
8	-5.584	13.11	19.82

9	-6.43	1.684	2.39

9	-4.655	4.352	8.341

9	-5.533	13.38	19.81

Contributions:

We worked together to create the research plan and report, and Aakriti focused on ABCB1 while Anusheh focused on CYP3A5.

Works Cited

Azam, Fahad, et al. “Influence of ABCB1 Gene Polymorphism on Concentration to Dose Ratio and Adverse Effects of Tacrolimus in Pakistani Liver Transplant Recipients.” *Pakistan Journal of Medical Sciences*, vol. 37, no. 3, May 2021, p. 689.

Brunet, Mercè, and Marçal Pastor-Anglada. “Insights into the Pharmacogenetics of Tacrolimus Pharmacokinetics and Pharmacodynamics.” *Pharmaceutics*, vol. 14, no. 9, Aug. 2022, p. 1755.

Dessilly, G., et al. “ABCB1 1199G>A Genetic Polymorphism (Rs2229109) Influences the Intracellular Accumulation of Tacrolimus in HEK293 and K562 Recombinant Cell Lines.” *PloS One*, vol. 9, no. 3, Mar. 2014, https://doi.org/10.1371/journal.pone.0091555.

Gutmann, Daniel A. P., et al. “Understanding Polyspecificity of Multidrug ABC Transporters: Closing in on the Gaps in ABCB1.” *Trends in Biochemical Sciences*, vol. 35, no. 1, Oct. 2009, p. 36.

Haverals, Lien, et al. “Does the Tacrolimus Trough Level Adequately Predict Drug Exposure in Patients Requiring a High Tacrolimus Dose?” *Transplantation Direct*, vol. 9, no. 4, Mar. 2023, p. e1439.

Islam, Md Ataul, et al. “Identification of Potential Cytochrome P450 3A5 Inhibitors: An Extensive Virtual Screening through Molecular Docking, Negative Image-Based Screening, Machine Learning and Molecular Dynamics Simulation Studies.” *International Journal of Molecular Sciences*, vol. 23, no. 16, Aug. 2022, p. 9374.

Shi, Baojie, et al. “Genotype-Guided Model Significantly Improves Accuracy of Tacrolimus Initial Dosing after Liver Transplantation.” *EClinicalMedicine*, vol. 55, Jan. 2023, p. 101752.

UniProtKB/SwissProt Variant VAR_015003. https://web.expasy.org/variant_pages/VAR_015003.html. Accessed 4 Dec. 2024.

Woillard, Jean-Baptiste, et al. “A Donor and Recipient Candidate Gene Association Study of Allograft Loss in Renal Transplant Recipients Receiving a Tacrolimus-Based Regimen.” *American Journal of Transplantation : Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons*, vol. 18, no. 12, Dec. 2018, pp. 2905–13.