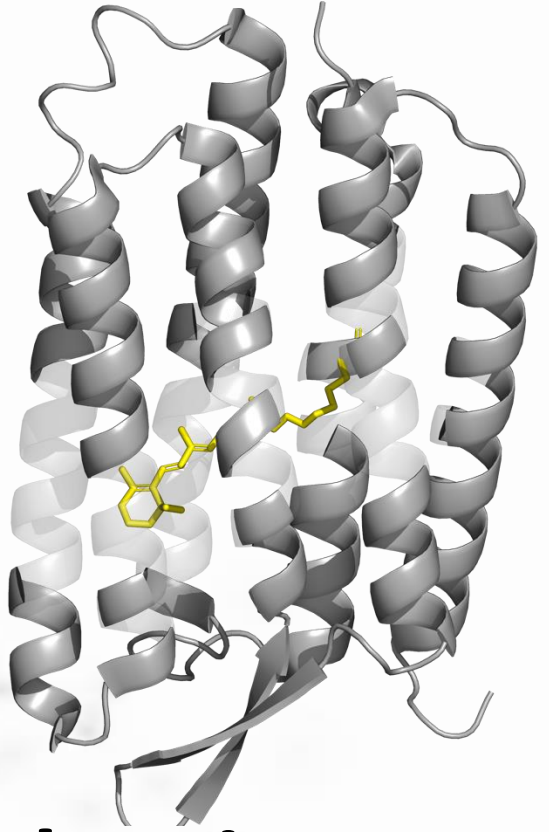
University of Idaho

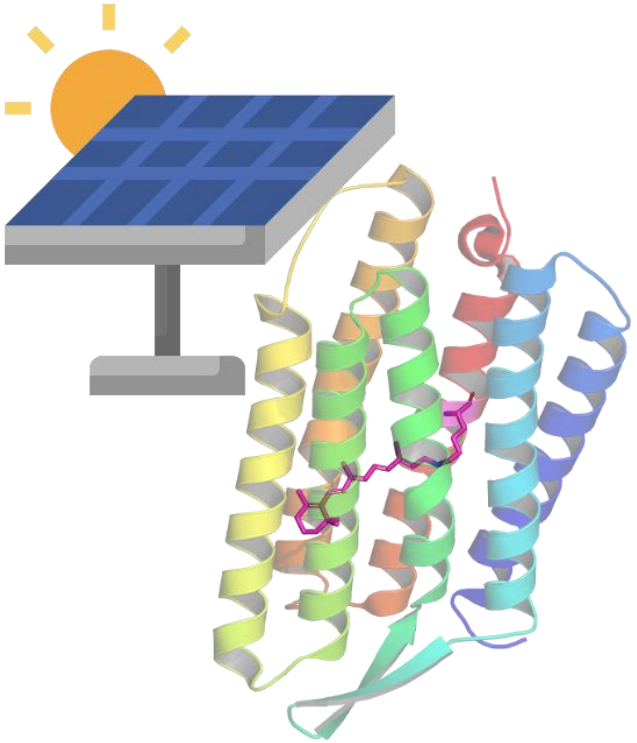
IMCI
Institute for Modeling Collaboration and Innovation



Unraveling Spectral Shifts in Microbial Rhodopsin through Machine Learning Predictions and Molecular Dynamics Simulations

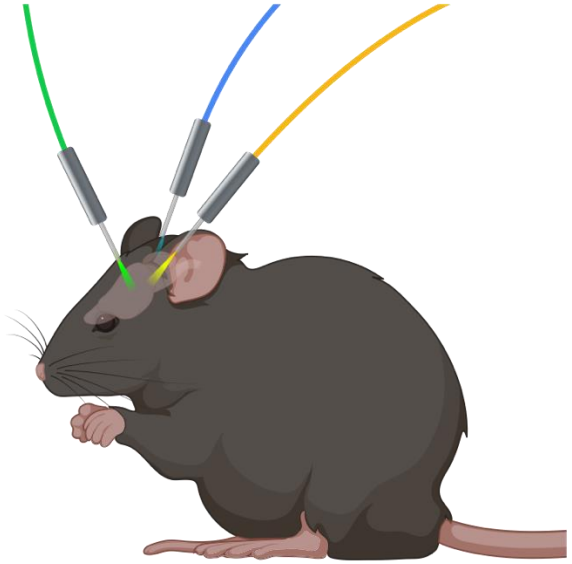
Shubham Pandey | PhD Student
Department of Chemical & Biological Engineering, University of Idaho

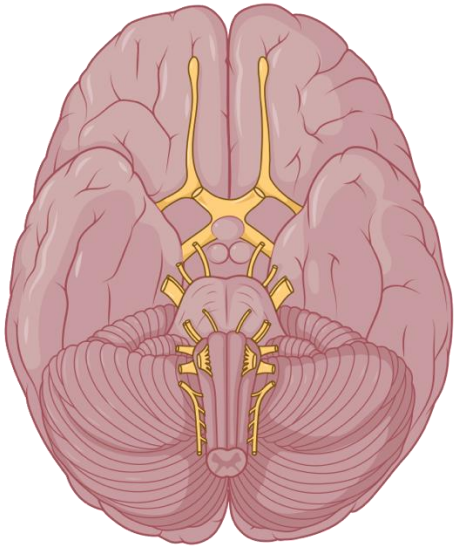
Life adapts to low-light environments through spectral tuning of Rhodopsin protein

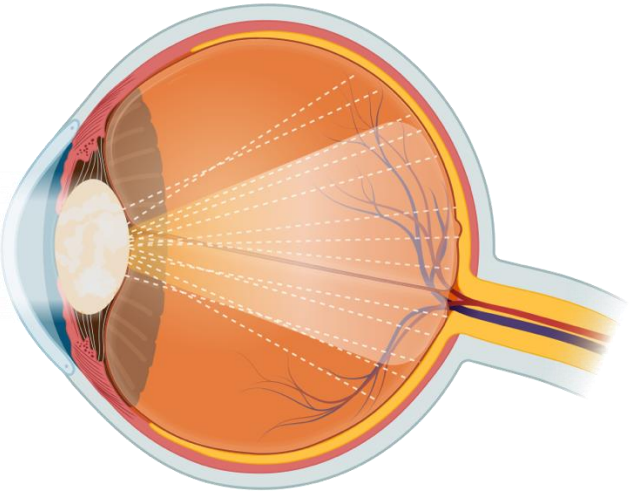


Generated using NightCafe AI.

Functional and Emerging Applications of Microbial Rhodopsins







Optogenetics

Light-activated control of ion transport and neuronal excitability

Neuroscience

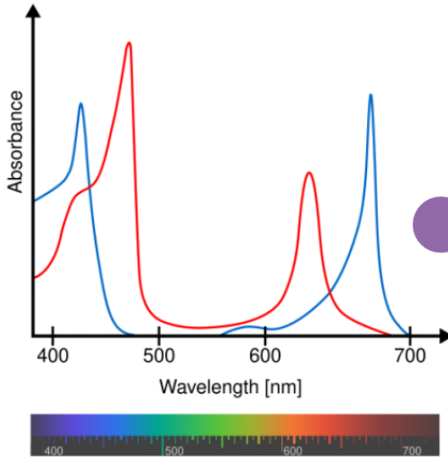
Functional mapping of neural circuits

Vision Science

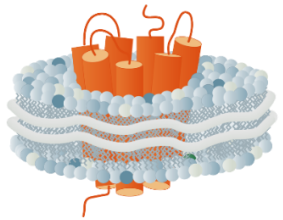
Restoration of photosensitivity in degenerated retinal cells

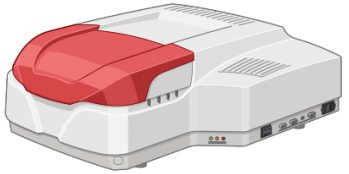
Accurate λ_{max} prediction enables rational spectral engineering, optimizes optogenetic tools, aids vision restoration, and reduces experimental cost and effort.

Explainable Machine Learning Framework for Sequence-Based λ_{max} Prediction in Microbial Rhodopsin

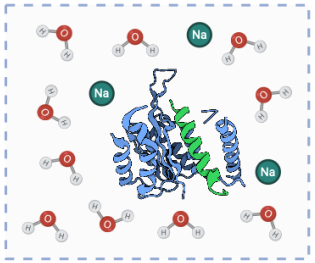


Experimental method





Membrane proteins are complex; experimental work is challenging, time-consuming, and laborious [± 2–5 nm error]



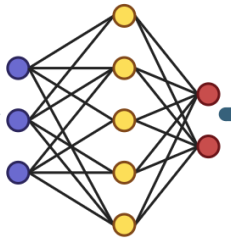
Computational method

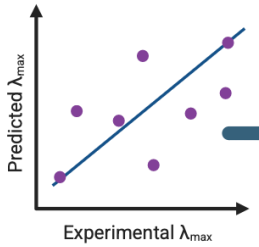
QM, QM/MM, MM

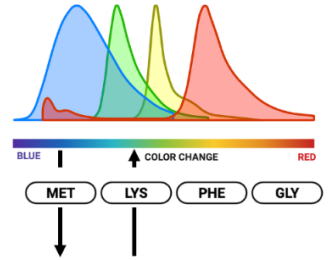
Computationally expensive, slow, and limited to small systems

Machine Learning Approaches

ML Prediction + SHAP Interpretation



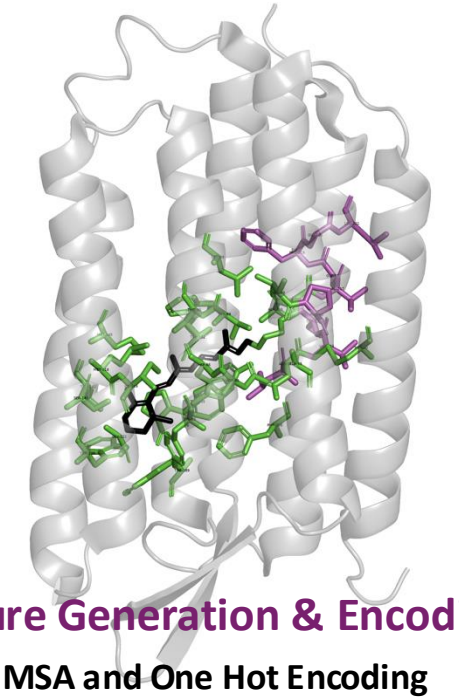




ML = Machine Learning

SHAP = Shapley Additive exPlanations

Machine Learning Pipeline for Predicting λ_{max}



Sequence Collection

884 rhodopsin sequences with experimental λ_{max} curated by Inoue *et al.*

PVTAEISKELFATEIEKYRESETN
TDVDNFDVDRFLVQKNFHYLPLDS
IRDLSGLSQKMVQTLLEQIRSNYD
YLTFSNTYTDEENETLINLEKTQS
LQKFMTQLDHLIKDDISNTQEIIK
VLEYLKKLDEIYGSLRNHSQLTEA
SLGKRLSKSLHEMCGIEPLEEEIC

Feature Generation & Encoding

MSA and One Hot Encoding

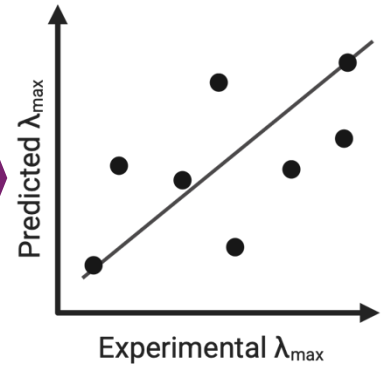
	Position 1	Position 2	… … …	Position 24
Seq 1	0	1	0	1	0	0	0	0	1	0
Seq 2	0	1	0	0	0	1	..	..	..	..
Seq 3	0	0	1	0	0	..	..	..	..	..
…										
	0	1	0	1	.	..	..	..	..	..
Seq 878	1	0	0	..	..	..	..	..	..	..

Model Training & Selection

Comparison of 13 machine learning regression models

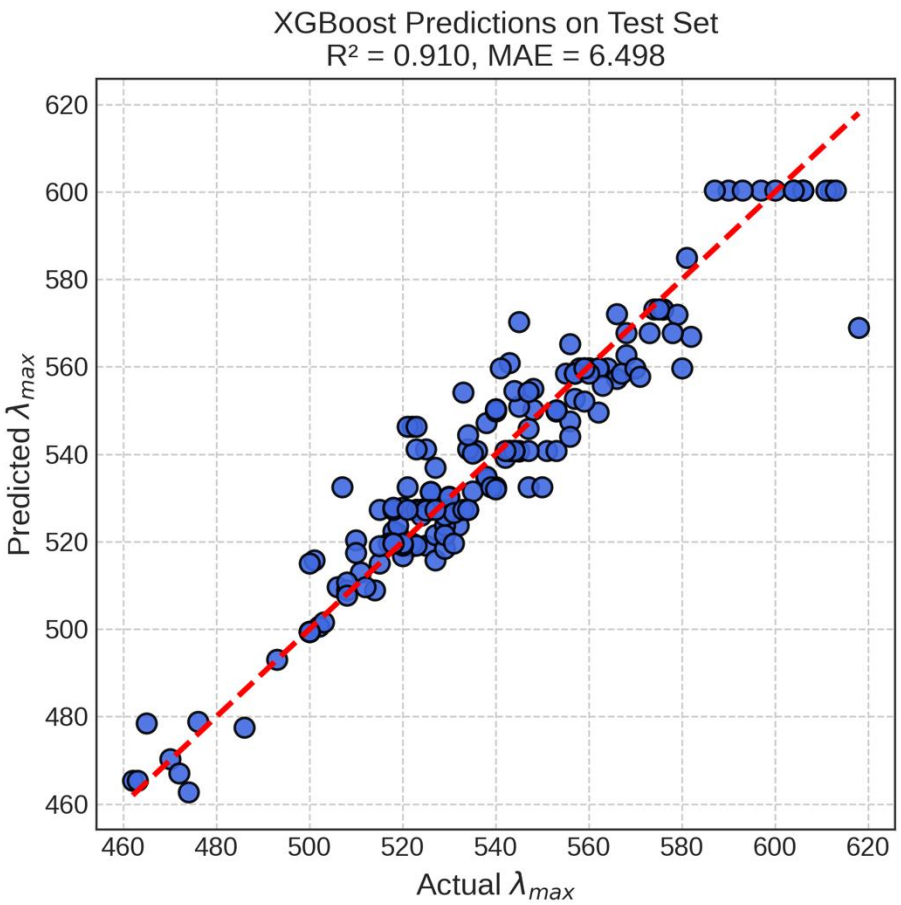
Prediction & Validation

Evaluate for max R^2 and min MAE

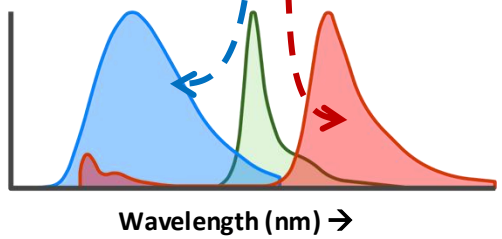


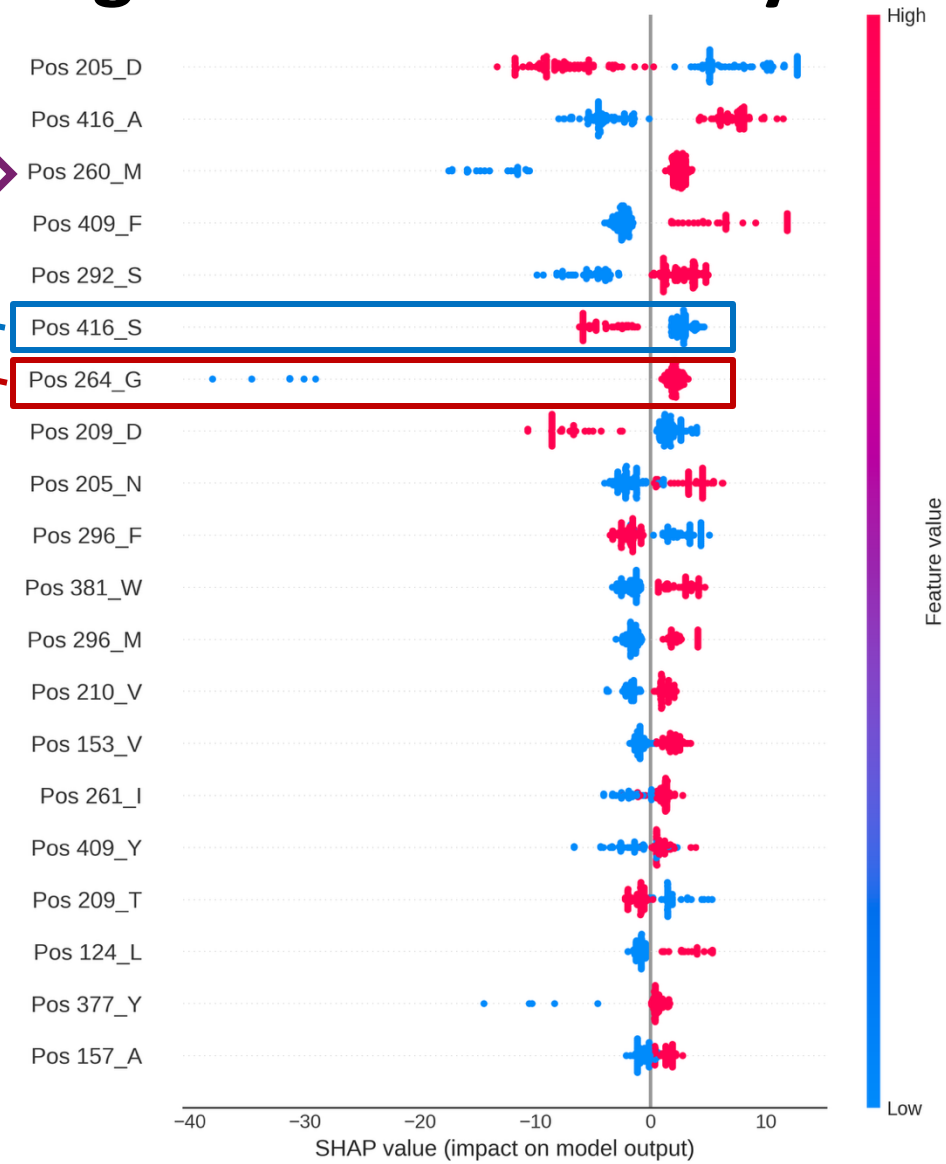
Machine Learning (ML) pipeline for λ_{max} prediction: feature generation, selection, model comparison, hyperparameter optimization, and performance evaluation.

Prediction Accuracy and Mechanistic Insights via SHAP Analysis



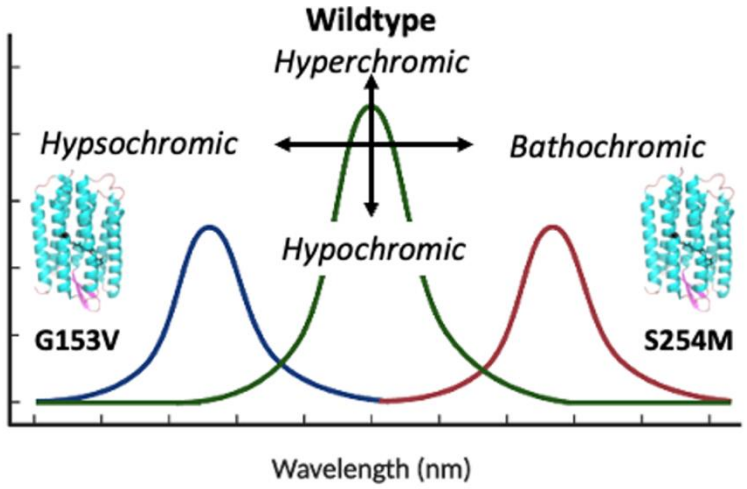
SHAP as an explanatory tool, interpreting XGBoost predictions

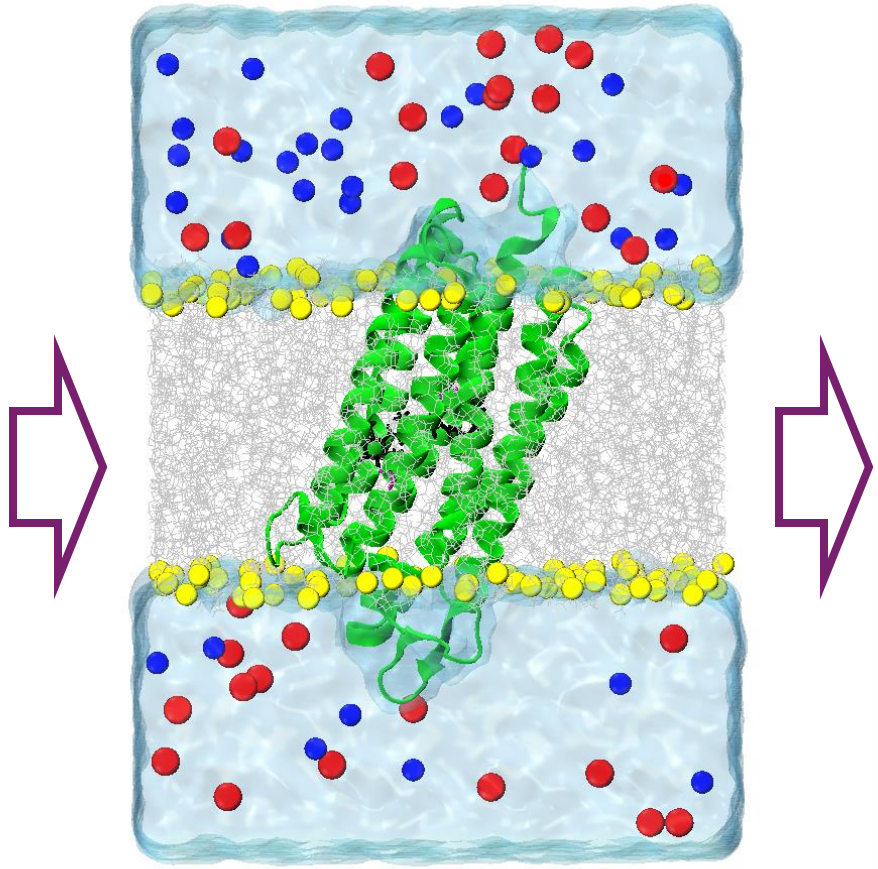


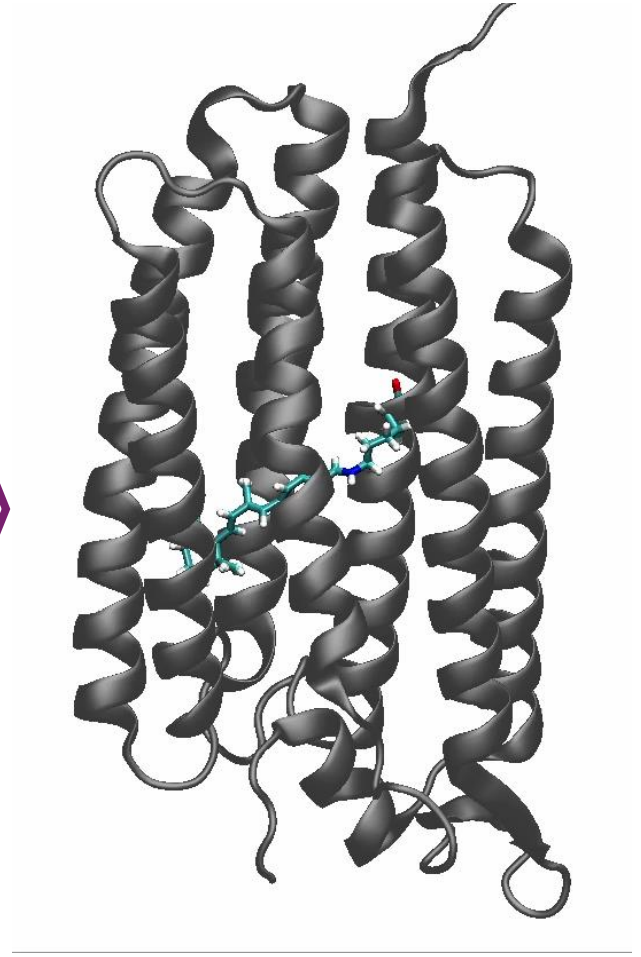


XGBoost accurately predicts rhodopsin λ_{max} (**MAE = 6.5 nm**, close to the **2–5 nm experimental uncertainty**)
SHAP analysis identifies key residues governing spectral tuning.

Validation of ML Predictions and SHAP Insights using Molecular Dynamics Simulation







Index of residues in Aligned file VS in KR2 (PDB: 3X3C)

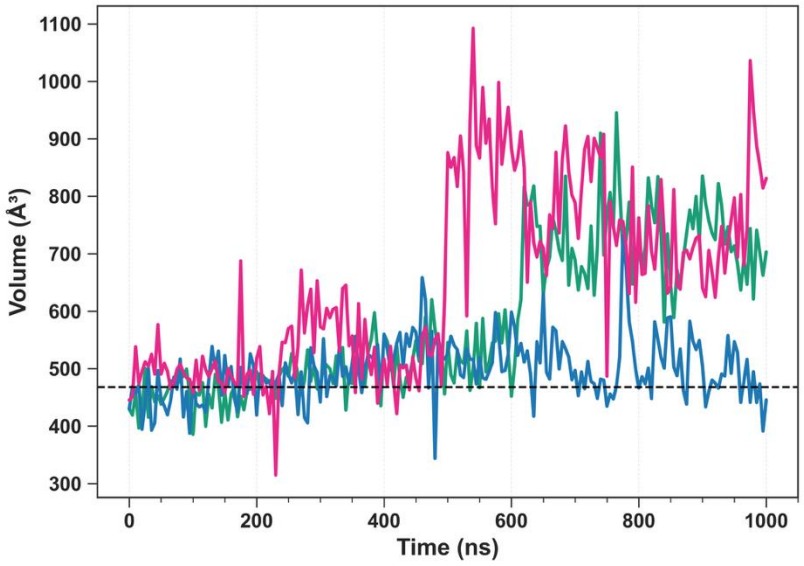
264 ←→ 153

416 ←→ 254

KR2 model used to verify ML predictions; MD captures mutation-induced retinal changes.

Molecular Insights from MD Simulation Reveal Mutation-Driven Spectral Shifts in KR2

Binding Pocket Remodeling



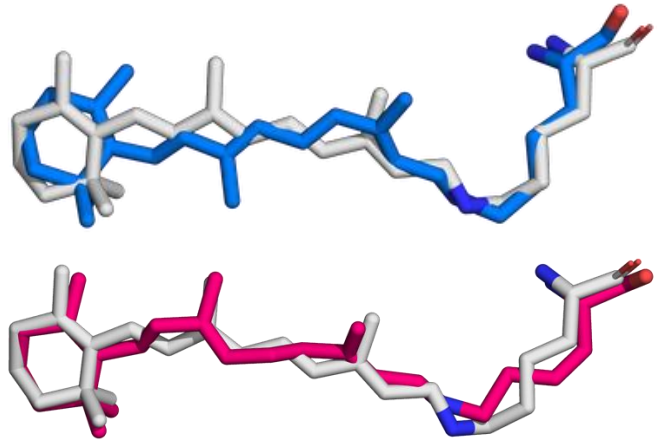
Retinal (Ref. 3X3C Coord.)

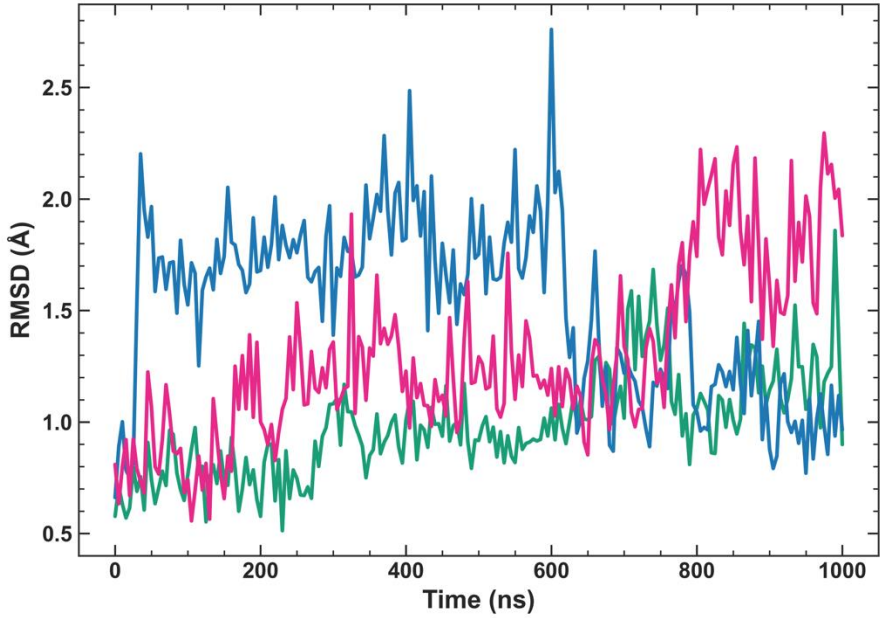
Wild Type

G153V

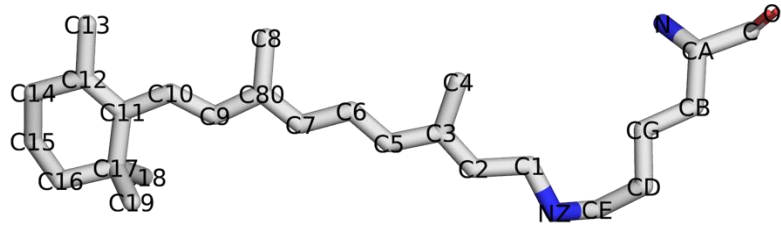
S254M

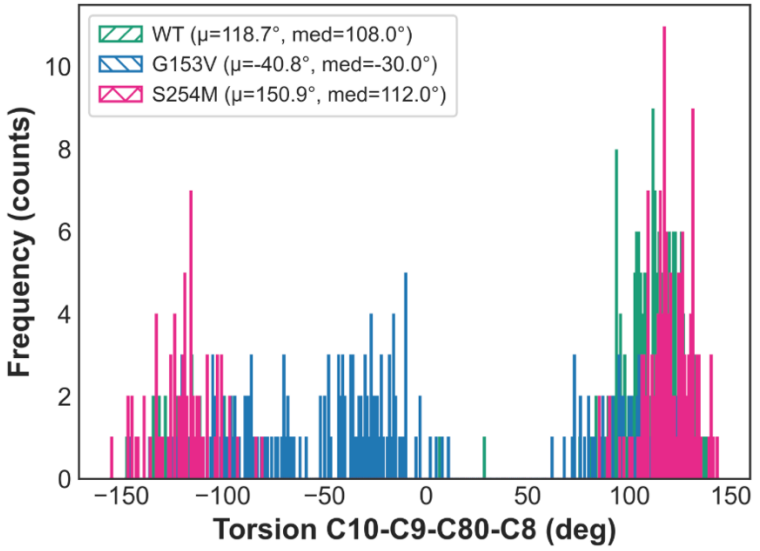
Retinal Conformation Change





Measure of Retinal Planarity





Results when correlated with Quantum Mechanical (QM) Principles

	S254M	G153V
Pocket Volume	Increase	Decrease
Retinal Planarity	More planar	More twisted/bent
Conjugation Length	Extended conjugation	Shortened conjugation
π-Electron Delocalization	Greater delocalization	Localized / impaired delocalization
	Red Shift (longer λ_{max})	**Blue Shift (shorter λ_{max})**

Acknowledgment

Dr. Jagdish Patel
(Advisor)

Research Group, Institutional Computational Support, and Funding Source

RESEARCH COMPUTING + DATA SERVICES

PatelDryLab

NIH National Institutes of Health

Open Source Machine Learning & Explainability Tools
Molecular Modeling & Simulation Developer Communities

scikit learn

Shap

CHARMM-GUI

FAST. FLEXIBLE. FREE. GROMACS

MD ANALYSIS

Thank You!