First Steps in Photoenzyme Design

PIs: **Thomas J. Lane**, Ron Dror, Henry van den Bedem, Possu Huang, Todd J. Martinez

SLAC NATIONAL ACCELERATOR LABORATORY

LELAND STANFORD JUNIOR UNIVERSITY

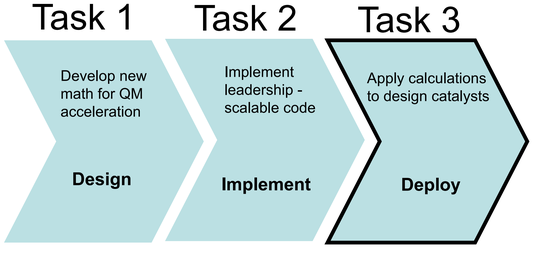
Motivation

Nearly all energy we use as a society originates from sunlight translated into chemical fuel by enzymes. This includes our food, fossil fuels, and biofuels. We are therefore motivated to study of any protein system that uses light to do chemistry.

We are out to answer the following questions:

- What are fundamental physics behind the molecular-scale modulation of photochemistry?
- What are the principles of operation behind existing photoenzyme systems?
- What computational tools can enable us to perform rapid and sophisticated protein design?
- How would we leverage answers to these questions to build a useful catalyst for a grand-challenge chemical problem: photoreduction of CO_2?

Our Role in the Collaboration



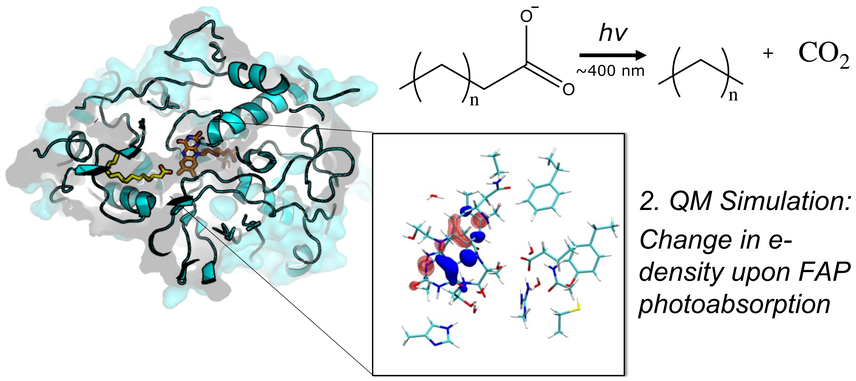
Designing Photocatalysis Through Scalable Quantum Mechanics

Raphael R. Eguchi & Po-Ssu Huang Automated Classification of Protein Structures by Semantic Segmentation. *Sci. Reports*. Submitted.

Raphael Townshend, Rishi Bedi & Ron Dror. Generalizable Protein Interface Prediction with End-to-End Learning. *NIPS 2018*. Submitted.

Sorigué et al., Science 357, 903–907 (2017)

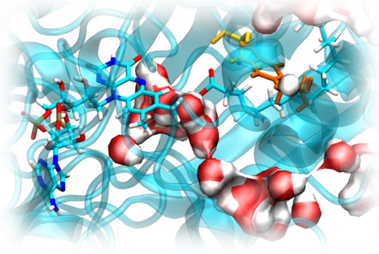
Wild Type System: FAP



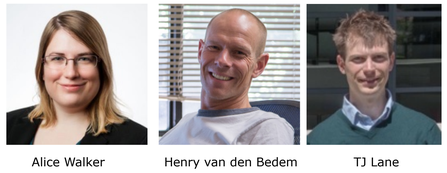
2. QM Simulation: Change in e-density upon FAP photoabsorption

Fatty Acid Photodecarboxylase (FAP)

- Photo-activated production of alkanes
- Requirement for light: not yet clear
- Mechanism: not known
- **Project goal**: QM calculations to predict mechanism, verify by comparing to spectroscopy, crystallography

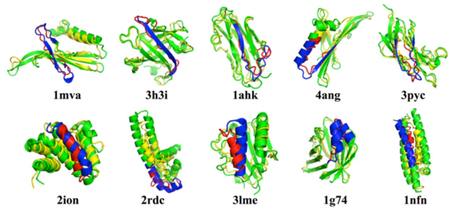


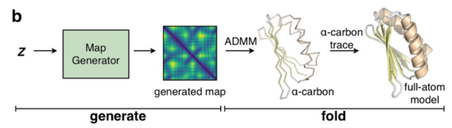
2. MD simulations predict water positions in low-resolution crystal structure active site. Long-timescale MD is being performed to predict substrate binding/ product release mechanisms, currently experimentally inaccessible.



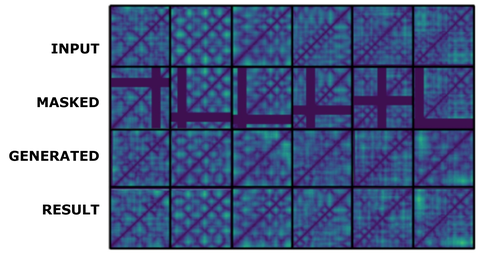
Alice Walker | Henry van den Bedem | TJ Lane

Cofactor Binding



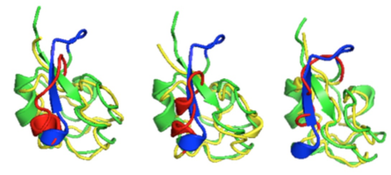


1. We have trained a GAN capable of "filling in" deleted sections of protein structures. The predicted structures closely match the original deleted segments and appear "protein-like".



Problem: Nearly all light-sensitive proteins employ an exogenous cofactor as a chromophore, but current methods for protein design cannot create cofactor binding sites.

Proposal: Employ generative models (NNs), trained on all protein structures, to create protein scaffolds consistent with complex restraints imposed by cofactor binding sites.



2. The generative approach allows us to create multiple designs that could be tested downstream.

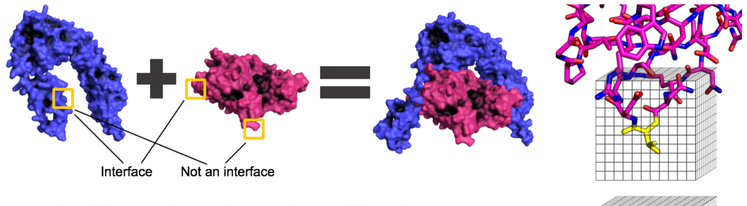
Possu Huang | Namrata Anad | Raphael Eguch

Submitted to *Scientific Reports*

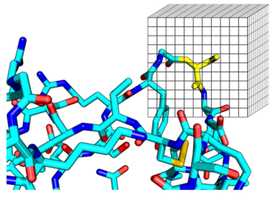
Protein-Protein Interfaces

Problem: Wild-type carboxylases exist (*e.g.* CODH), but need to be coupled to a source of high-voltage electrons. Design of protein-protein interfaces is not yet possible.

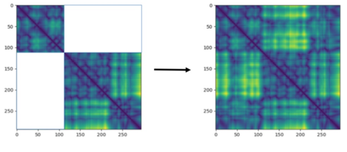
Proposal: Learn how to design protein-protein interfaces to bind a chosen (perhaps designed) oxidase to wild-type reductase for catalysis.



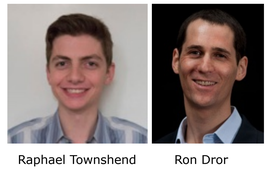
1. Even interface identification is challenging, as structures often deform upon binding and interactions are delocalized.



2. We have trained a NN that predicts PPI to accuracies (0.93) better than state of the art (0.90).



3. Design of PPI can be thought of an extension of cofactor binding (above).



Raphael Townshend | Ron Dror

Submitted to NIPS