

# Developing Computational Methods for Enzyme Design

Hasti Sanandajifar, Scott Johnson, K.N. Houk  
University of California, Los Angeles

## Abstract

As part of an ongoing project to develop methods for the computational design of enzymes, docking programs were evaluated based on three criteria: 1) Does the docking program predict binding of the correct substrate? 2) Does the docking program correctly predict that the enzyme binds the transition state analog more strongly than it does the ground state? 3) Does the docking program predict that the enzyme binds the product less strongly than it does the substrate? The study focused on the performance of the Molegro Virtual Docker (MVD) program and the docking program FRED. The enzyme chorismate mutase, which catalyzes the reaction of chorismate into prephenate, and a variety of mutants were explored, since the quantitative experimental data about binding of substrate, transition state and product are known and 3D structures of the proteins are available. The interactions of chorismate mutase enzyme and various point mutations, with chorismate, a transition state analog, and prephenate were tested. Initial comparisons of docking results with available kinetic data display no correlation between docking scores and the presence or absence of bonds known to be necessary for catalysis versus the rate of catalysis or  $K_M$ . The respective binding affinities calculated for chorismate, a transition state analog inhibitor, and prephenate display the correct trend, but the differences between binding affinities are found to fall within the error range of the program for the prediction of binding affinities. Thus the programs calculate the magnitudes of docking correctly but not precisely enough to numerically differentiate between different poses.

## Introduction

Enzymes play essential roles in nature by catalyzing reactions in all organisms. But is it possible to build enzymes to catalyze specific reactions that do not occur in nature? The ultimate goal of our project is to create algorithms for the computational design and evaluation of enzymes for any given reaction. Approaches to this problem include the use of docking programs to assess how they predict enzyme-substrate and enzyme-product binding. This analysis was performed for known enzymes and will then be applied to enzymes that have not been experimentally tested.

A good enzyme will readily bind to a substrate but readily release the product. Finding ways to determine the strength of enzyme-substrate and enzyme-product binding with docking programs will save the costs and time involved in experimentally testing a large variety of proteins to find out their ability to catalyze given reactions. Several docking programs are Molegro, FRED, Glide, Dock6, MVP, and Autodock. By calculating binding energies, interatomic forces, as well as a variety of other factors and incorporating them into a scoring function, these docking programs computationally predict how well specific small molecules bind to a particular protein. Typically these programs are used to determine protein-ligand or drug-inhibitor binding, but we have been assessing their accuracy in predicting enzyme-substrate binding.

For a good enzyme design we need to have a small  $K_M$  for the substrate (to ensure strong binding) and a large  $K_P$  for the product (to allow product release). Research has shown that the presence of catalysis is dependent on the presence of particular bonds between a ligand and specific amino acid residues within an enzyme. We have been using docking to find energetically and sterically favorable poses for the enzyme-substrate interactions and comparing the trends in binding affinity to experimental results. We have also been looking at the poses to see whether the necessary bonds between enzyme and ligand are present. We have examined some of the docking parameters, the energies of different poses, and the RMSD (root mean square deviation) values of the poses as compared to the experimental pose to see how docking can most accurately predict enzyme-substrate interactions. This information can then be used in conjunction with transition state binding energies and known information about protein structure to help computationally design enzymes for specific reactions.

## Background

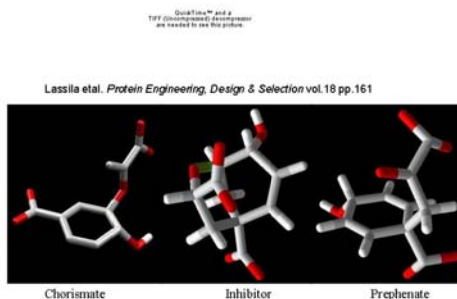
**Reaction Rates and Binding Affinity:**  
Experimentally, the rate of substrate going to product is  
rate =  $k_{cat} [S] [E] / (K_M + [S])$   
Where  $K_M = ([E][S]) / [ES]$

The binding affinities of the substrate and product to the enzyme are related to  $K_M$ , whereas the binding affinity of the transition state to the enzyme is related to the  $K_{TS}$ .  
In our case  $K_P \gg K_M$  because  $k_{cat} \ll k_{-1}$

**Calculation of binding affinities by Molegro:**  
 $E_{binding} = -5.68 \cdot \log K_i$  (kJ/mol) where  $T = 297K$   
 $K_i$  estimated using "a combinations of energy terms and molecular descriptors"

\* For 200 test structures with known binding affinities, molegro gave results with a correlation coefficient of 0.60 (MVD user manual)

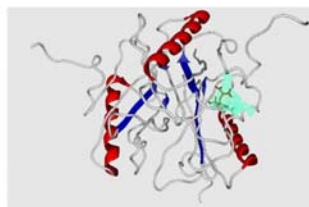
**Chorismate Mutase:**  
catalyzes the Claisen rearrangement of chorismate to prephenate.



(Chorismate and prephenate obtained by changing connectivities in structure of inhibitor and minimizing the resulting structure.)

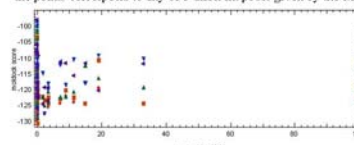
## Overview of Docking

Docking programs evaluate binding of a large molecule, such as a protein, with a small ligand, such as our substrate. For various poses calculated to be energetically and sterically most favorable, they usually give an overall score, binding affinity, hydrogen bonding score, and root mean square deviation (RMSD) from the position of the original substrate.

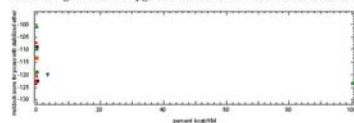


## Results

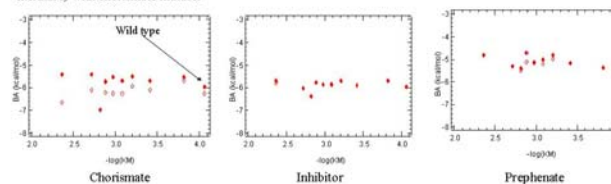
Does the docking score for chorismate (the substrate) have any correlation with the catalytic rate? The following graph shows that there is no correlation between score (calculated by Molegro, using PLP scoring function) and percent catalytic activity. (The different colors for the points correspond to any of 5 different poses given by the Molegro docking program.)



Does the docking score in conjunction with the presence/absence of bonds predict catalytic activity? The following plots the docking scores only of the poses with the arg90 residue stabilizing the ether oxygen in chorismate. \* No conclusion can be drawn from these results.



Although the binding affinity of the substrate to chorismate mutase did not correlate with the expected  $K_M$ , some potentially interesting results were observed when binding affinities of chorismate, its transition state analog, and prephenate were compared. In the following plots the unfilled diamonds represent the highest binding affinity given for docking of chorismate in a particular mutant of chorismate mutase, and the filled diamonds represent the binding affinity of the pose making the most catalytic bonds (as obtained from literature) with chorismate mutase.



The plots show that prephenate, the product, binds less strongly to chorismate mutase than chorismate and the inhibitor do. However, due to current program limitations binding affinities are predicted to within 2 kcal/mol (and the observable difference between prephenate and the inhibitor is about 1 kcal/mol); therefore there is no conclusive evidence that the program does in fact predict that prephenate binds less strongly.

## Comparison with Crystal Structure:

The crystal structure for the inhibitor inside of wild-type chorismate mutase is available. Comparing the best pose given by Molegro for the wild type with the actual crystal structure shows that molegro can give the correct pose, even if this pose is not ranked first in score or binding affinity. In the picture to the right, the inhibitor and its pose are shown, with the docked pose highlighted in green.

## Comparison with FRED docking program:

Because FRED does not calculate binding affinities we used Molegro to calculate the binding affinities of the various poses given by FRED. Chorismate was docked using a consensus scoring function including PLP, Chemscore3, and oeChemscore to find the best pose.

Comparing the crystal structure of the inhibitor inside chorismate mutase with the poses given by FRED gave similarly good results. FRED gave a smaller variation between different poses. The picture at the right shows the best pose given by FRED (highlighted) with the protein and inhibitor crystal structure.

## Future Goals

- Calibrate the Molegro docking program by comparing results to other programs.
- See if crystal structures could be found for point mutations of chorismate mutase.
- Dock the transition state of the chorismate reaction.
- Run REDCat (if available) on the docked structures
  - REDCat is a program that ranks poses of a ligand based on the presence of catalytic bonds between the ligand and particular amino acid residues in the protein.
- Dock other enzymes to see if these results are reproducible

## References

1. Lesslila, J; Keffe, J; Oelschlaeger, P; Mayo, S.: "Computationally designed variants of Escherichia coli chorismate mutase show altered catalytic activity." *Protein Engineering, Design, and Selection* 2005, 18, 161-163.
2. Houk, K; Leach, A; Kim, S; Zhang, X.: "Binding Affinities of Host-Guest, Protein-Ligand, and Protein-Transition-State Complexes." *Angewandte Chem. Int. Ed.* 2003, 42, 4872-4897.
3. Warren, G. et al.: "A Critical Assessment of Docking Programs and Scoring Functions." *Journal of Medical Chemistry* 2006, 49, 5912-5931.