Binding Interactions between Plant-Based Phenolic Chemicals and Horseradish Peroxidase

Jesica R. Williams, PhD1, Willie F. Harper, Jr., PhD, PE2

1Oak Ridge Institute for Science and Education Research Participation Program, Environmental Engineering and Science Program, Air Force Institute of Technology, Dayton, OH, USA
2Department of Systems Engineering and Management, Environmental Engineering and Science Program, Air Force Institute of Technology, Dayton, OH, USA

Abstract—Horseradish peroxidase (HRP) has the potential to help remediate environmental settings that are contaminated with an array of persistent pollutants. This study investigated chemical interactions between HRP and plant-based phenolic chemicals (PBPCs). The average binding distance between the substrate phenolic hydrogen and the imidazole delta nitrogen (δN) of HIS42 was 8.46 Å but this binding distance was poorly correlated to binding energy (ranging between -4.9 and -7.5 kcal/mol) for the wild type complexes. Mutant models demonstrated that electrostatic interactions between the carboxylic group of the substrate and ARG38 compensated for unexpectedly weaker hydrogen bonding interactions between the phenolic hydrogen of the substrate HIS42. HRP-PBPC binding interactions are often stabilized by non-active site interactions with ASP150 and SER151. An artificial neural network (ANN) model successfully related substrate chemical properties with binding energy when 80% of the available data set is used for training. This work is the first to provide computational analyses on HRP interactions with PBPCs and the results are expected to inform future research approaches and clean-up efforts.

Keywords—Horseradish peroxidise, binding energy, molecular docking, chemical interactions.

I. INTRODUCTION

Water can be treated with horseradish peroxidase (HRP), a heme-containing enzyme that uses hydrogen peroxide and a phenolic substrate to generate highly reactive hydroxyl radicals [1,2]. HRP can be attached to a substrate or suspended in solution and then activated to transform a wide variety of chemical pollutants including pharmaceuticals [1,3], chlorinated organics [4], industrial dyes [5], and oils [6]. The versatility of this enzyme is further underscored by its stability over a wide range of pH and temperature [7,8].

HRP has the potential to flexibly support ongoing efforts to protect water quality and remediate settings that have been contaminated with an array of persistent pollutants with varying chemical properties. The most abundant HRP form is Isoenzyme C (HRP C), available from the root of the horseradish plant. HRP C contains 308 amino acid residues and it contains both alpha and beta sheet structural regions with the heme group located between distal and proximal domains [9]. Structural studies have produced a 3D crystallographic model at 1.6Å resolution and a crystallographic R-factor of 0.188 [10,11]. The active site residues are ARG38, PHE41, HIS42, ASN70; ARG38 stabilizes the substrate with electrostatic interactions, PHE41 prevents the substrate from accessing the ferryl oxygen, HIS42 forms a hydrogen bond with the phenolic proton of the substrate, and ASN70 maintains the basicity of the HIS42 residue [9]. Colosi et al., 2006 previously studied HRP activity against a panel of phenolic estrogens and suggested that the rate-controlling chemical interaction was between the phenolic hydrogen of the substrate and the imidazole δN of HIS42 [1]. This insight was largely revealed by using both molecular simulations and site-directed mutagenesis; the former can inform experimental efforts while the later provides validation.

HRP directly reacts with phenolic chemicals, and normally, phenol or another similar compound is added along with hydrogen peroxide. Another alternative is the use of plant-based phenolic compounds (PBPCs), which are naturally occurring and have not yet been leveraged as a substrate for HRP. PBPCs are found in the seeds, leaves, and flowers of plants [12], and they afford protection against pathogens, parasites, and oxidative stress [13]. PBPCs may offer a new and more economical source for a phenolic substrate, but these chemicals are considerably larger than commonly-used phenol and it is not clear that the binding properties discovered to date would apply when the reactions are driven by PBPCs.
The overall objective of the current work is to computationally simulate PBPC interactions with HRP. Molecular docking simulations were performed to relate substrate binding distance with binding energy, which is well-known to correlate with enzymatic kinetics [14]. Wide type (WT) and mutant HRP model complexes were developed to assess the involvement of additional active site residues. An artificial neural network (ANN) model was also developed to assist with the development of structure-activity relationships.

II. MATERIALS AND METHODS

A. Enzyme-Substrate Docking

Each of thirty PBPCs was flexibly docked into a rigid model of the HRP enzyme using the AutoDock Vina software program [15]. Vina uses an Iterated Local Search global optimizer and the scoring function ranks docked conformations based on the sum of intermolecular and intramolecular contributions [15,16]. The advantage of AutoDock Vina is it improves the speed of docking from parallelism by multithreading on multicore machines. Vina also achieves higher substrate binding mode accuracy as it is equipped to deal with more active rotatable bonds and automatically calculates its own grid maps and clusters the results [15].

The HRP enzyme structure was downloaded from the Protein Data Bank (PDB: 1H55) and both the enzyme and substrates were prepared for docking using AutoDock Tools [17]. Default docking parameters were used to generate each enzyme-substrate complex: binding modes = 9, exhaustiveness of search = 8, and maximum energy difference = 3. AutoDock Vina uses an iterative local search global optimizer for the optimization algorithm [17,18,19] and the clusters of docked conformations were assessed. The lowest binding energy (kcal/mol) pose generated for each complex was selected for further analysis. The enzyme-substrate binding interaction distances for each complex were computed by measuring the distance in angstroms between the phenolic proton of each substrate and the delta nitrogen (δN) of the imidazole ring of the HIS42 using the Pymol software program [20].

B. HRP Mutant Models

HRP mutant models were developed using Pymol based on previously published mutagenesis experiments [21,22]. Phenylalanine 41 was mutated to alanine, leucine, and threonine residues (PHE41A/L/T) and the thirty plant-based phenolic substrates were docked into the mutant models using AutoDock Vina to investigate HRP mutant enzyme-substrate interactions. The ALA41, LEU41, and THR41 mutants are smaller in size relative to PHE41 and were predicted to keep the HRP active site open for substrate binding. Published data shows the PHE41A/L/T mutations to both increase the rate of HRP catalysis as well as decrease the average binding distance between the phenolic proton of the substrate and HIS42.

C. Quantitative Structure-Activity Relationship Development

Two-dimensional QSAR profiles were developed to relate substrate binding energies with substrate properties. Substrate average mass, the number of hydrogen bond acceptors, the number of hydrogen bond donors, and the distance between the phenolic proton of the substrate and the imidazole delta nitrogen of HIS42 were the substrate properties compared to the enzyme-substrate binding energies. Each PBPC was docked into the HRP WT and mutant models (PHE41A/L/T). Each docked model complex was assessed for binding energies and binding distances between the substrate and the catalytic resides ARG38, PHE41, HIS42, and ASN70.

D. Artificial Neural Network (ANN)

An artificial-neural network (ANN) model was developed from automated docking simulations between the thirty plant-based phenolic substrates and the horseradish-peroxidase enzyme. The input matrix contained substrate properties and the output matrix contained enzyme-substrate binding energies. Optimum results were observed when 80% of the data was used for training, 10% of the data was used for validation, and 10% of the data was used for testing. The Levenberg-Marquardt back propagation algorithm was employed during network training. The performance of the ANN model was estimated during the validation phase and the overall performance of the model was evaluated in the testing phase.

III. RESULTS AND DISCUSSION

E. The Effect of Substrate Properties on Binding Energy

Panel A of Figure 1 shows the correlation between enzyme-substrate binding energies and substrate atomic masses.
F. The Effect of HRP Mutations on Binding Properties

Mutation has the potential to increase the rate of HRP catalysis by shortening the average binding distance between the phenolic hydrogen of the substrate and HIS42. Replacement with a non-polar aromatic sidechain of phenylalanine (F) would add a larger chemical group, relative to the alanine (A), leucine (L), and threonine (T) mutants while mutation to the smaller non-polar hydrocarbon sidechains of ALA and LEU and the polar side chain of THR may result in increased enzyme catalysis likely due to a more open active site. For example, Figure 2 is representative of the docked model complex of Caffeic Acid with HRP WT; the binding energy for this interaction is -5.8 kcal/mol, the distance between the HIS42 and the phenolic proton of the substrate is 7.6 Å, and the distance from the carboxylate oxygen to the ARG38 residue is 3.6 Å. When Caffeic Acid is docked with HRP WT and mutants, the calculated binding energies are similar and between -5.7 and -6 kcal/mol. Thus, the mutations did not greatly impact the binding energy. These results further suggest that the HIS42 binding distance does not solely drive the values of the binding energies for HRP-PBPC interactions.

Figure 3 shows the molecular docking simulation of HRP WT and 3-Hydroxybenzoic Acid with a binding energy of -5.0 kcal/mol. Active site residues are colored teal and the substrate is blue. The carboxylate group binds 3.4 Å from ARG38 and the phenolic hydrogen of the substrate binds 7.9 Å from HIS42. Mutation of PHE41 to ALA41, a smaller residue, is expected to open the active site for substrate binding resulting in more favorable binding energies, stronger ionic/electrostatic interactions between ARG38 and the carboxylate group of the substrate, and stronger hydrogen bonding interactions between the phenolic proton of the substrate with HIS42. The model indicates that this mutation causes more favorable (i.e. more negative) binding energies. However stronger hydrogen bonding interactions with HIS42 were not observed, which also points toward its limited effect on the binding energies. Substrates had weaker interactions with HIS42, shown by a longer binding distance of 8.1 Å for F41A, 8.8 Å for F41L, and 11.6 Å for F41T. Interestingly the F41 mutants resulted in stronger ionic/electrostatic interactions between ARG38 and the carboxylate group of the substrate as shown by a shorter binding distances of 3.1 Å for F41A, 3.1 Å for F41L, and 3.0 Å for F41T. These results point to the ARG38 binding distance as a key metric in F41 mutants, which should be investigated further using site directed mutagenesis.
Figure 1. Binding energy vs. substrate properties. Panel A is a plot of binding energy vs. binding distance. The distances in angstroms between the phenolic proton of each substrate and the δN of the imidazole ring of HIS42 (µHIS42) are plotted on the x-axis. Panel B is a plot of substrate binding energy vs. substrate average mass. Panel C is a plot of substrate binding energy vs. the number of hydrogen bond acceptors. Panel D is a plot of substrate binding energy vs. the number of hydrogen bond donors.

Figure 2. Docked model complex of Caffeic Acid and HRP WT. Caffeic acid shown in sticks colored light blue is predicted to bind to HRP WT with a binding energy of -5.8 kcal/mol. Active site residues are cyan; the heme group is colored dark blue.

Figure 3. Docked model complex of 3-Hydroxybenzoic Acid and HRP WT. HRP is colored magenta and active site residues are colored teal. The substrate is colored dark blue and the binding distances to ARG38 and HIS42 are 3.4Å and 7.9Å, respectively.
Docking of HRP WT with Salicylic Acid revealed a binding energy of -5.0kcal/mol (not shown). Binding interaction distances observed between the substrate and enzyme were greater than 10 Å; this long distance suggests the substrate will not be a strong activator of HRP. The F41A/L/T mutants were expected to open the active site for more favorable binding energies due to stronger ionic/electrostatic interactions between the carboxylate group of the substrate with ARG38 and stronger hydrogen bonding interactions between the phenolic proton of the substrate with HIS42. However weaker interactions were observed as shown with greater than 10 Å binding distances between the substrate and the mutants. As expected, lower binding energies were observed at -5.2kcal/mol, -5.2kcal/mol, and -5.6 kcal/mol respectively for the F41A/L/T mutants. The lower binding energies of the mutants suggest the enzyme-substrate complex is stabilized by non-active site interactions with ASP150 and SER151 at 4.6 and 6.3 Å, respectively.

G. Addressing Complexity

The relationship between substrate properties and binding energy is complex and multifaceted, as demonstrated by the current results. Therefore, an artificial-neural network (ANN) model was developed to relate the data sets (Figure 4). The input matrix included the phenolic average mass, the number of hydrogen bond acceptors, the number of hydrogen bond donors, and the binding distance to the HIS42 residue and the output matrix contained binding energy. When 80% of the available data set is used for training, a strong correlation (i.e. R² =1) is observed. ANNs are generally employed when complex relationships are to be mapped (reference), and this result shows that ANNs can be used to assist with the screening of PBPCs for HRP-driven reactions.

Figure 4 also maps numerous examples of structurally different chemicals with identical binding energies. For example, Isorhamnetin and Chlorogenic Acid both have a binding energy of -6.7kcal/mol. In this case, the C1 phenolic hydrogen of Isorhamnetin (average mass of 316.262 Da) binds 7.6 Å from the imidazole δnitrogen of HIS42 while the C1 phenolic hydrogen of Chlorogenic Acid (average mass of 354.309 Da) binds 7.5 Å from the imidazole δN of HIS42.
The model predicts the similar binding interaction distances of Isorhamnetin and Chlorogenic Acid, at 7.6 and 7.5 angstroms, contributes to the common binding energy of -6.7 kcal/mol. However, there are examples of structurally different phenolics that are associated with the same binding energy, but with very different chemical interactions at play. For example, resveratrol and transacin are both associated with a binding energy of -6.4 kcal/mol.

In both cases there are relatively weak interactions between the phenolic hydrogen and the HIS42 (i.e. the binding distances to the HIS42 are 5.9 Å and 7.8 Å for resveratrol and transacin respectively). The weak HIS42 interaction is compensated for in different ways. In the case of resveratrol there are pi-pi stacking interactions between the aromatic sidechains of PHE68 and Resveratrol at 4.4 Å, but in the case of Transacin, there are hydrogen bonding interactions between the phenolic hydrogen and the amide oxygen of PRO139 at 2.8 Å.

The central heme group of HRP sterically hinders the hydrocarbon tail of transacin from binding deeper in the active site, thus preventing stronger interactions. Identical binding energies can result from the resultant sum of different chemical forces.

Another factor adding complexity arises from the fact that enzymatic activity, while proportional to binding energy, may not be directly related to the observed degradation rates of particular chemical pollutants. The reason is the nature of the hydroxyl radical, which can attack organic pollutants in one of three oxidative modes: 1) hydrogen abstraction (i.e. removing a hydrogen atom from a saturated hydrocarbon), 2) hydroxylation (i.e. adding the hydroxyl group to an unsaturated hydrocarbon), or 3) oxidation without transfer of atoms. The nature of the chemical pollutant of interest will determine the relevant oxidative mechanism, which in turn will impact the shape of the pre-reactive (i.e. transition state) complexes.
Hydrogen abstraction reactions occur on saturated hydrocarbons, while hydroxylation occurs at unsaturated hydrocarbon bonds [24]. For example, aromatic compounds with nonpolar side chains (e.g. phenylalanine) are likely to be subject to hydrogen abstraction along the side chain but hydroxylation on the aromatic ring. Hydroxylation is likely to be a dominant mechanism during initiating reactions for many persistent organic pollutants because of the presence of numerous C-C double bonds, however, during the ensuing byproduct reactions the transformations will depend on the structure of the metabolites. Developing relationships between phenolic properties and observed pollutant degradation rates will require coupling HRP-related QSARs with expressions that describe interactions between hydroxyl radicals and chemical pollutants.

IV. CONCLUSIONS

This work is the first to provide computational analyses on HRP interactions with plant-based phenolic substrates (PBPCs): (1) QSAR models suggests binding energy is not directly related to binding distance for HRP-PBPC interactions; (2) Molecular docking simulations suggest electrostatic interactions between the carboxylic group of the substrate and ARG38 compensates for weaker hydrogen bonding interactions between the phenolic hydrogen of the substrate HIS42; (3) HRP-PBPC binding interactions are stabilized by non-active site interactions with ASP150 and SER151; (4) ANNs correctly predict HRP-substrate binding energies when trained with a 80% of available data. The computational models presented in the current work will guide future kinetic experiments to determine the effect of PBPCs substrates on the HRP activity.

V. ACKNOWLEDGEMENTS

The authors acknowledge the support of the Defense Environmental Restoration Account program.

VI. DISCLAIMER

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Air Force Institute of Technology, United States Air Force, Department of Defense, or United States government.

REFERENCES


