

Validation of a Two-Parameter Quantitative Structure–Activity Relationship as a Legitimate Tool for Rational Re-Design of Horseradish Peroxidase

Lisa M. Colosi,¹ Qingguo Huang,^{2,3} Walter J. Weber Jr.²

¹Department of Civil & Environmental Engineering, University of Michigan, Ann Arbor, Michigan

²Department of Chemical Engineering, University of Michigan, 2200 Bonisteel Blvd, 4103 ERB I, Ann Arbor, Michigan 48109-2099; telephone: +1-734-763-1464; fax: +1-734-936-4391; e-mail: wjwjr@umich.edu

³Department of Crop and Soil Sciences, University of Georgia, Griffin, Georgia

Received 17 November 2006; accepted 26 February 2007

Published online 9 March 2007 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.21419

ABSTRACT: Previously reported rates of reaction between six mutant strains of the enzyme horseradish peroxidase (HRP) and a test substrate, 2-methoxyphenol, were found to correlate with characteristic binding distances computed using molecular simulation. The correlation ($R^2 = 0.86$) bears out a working hypothesis that, based on a quantitative structure–activity relationship (QSAR) we had previously developed for HRP, reductions in binding distances between the HRP enzyme and any selected substrate mediate increased enzyme reactivity towards that substrate. The results validate the use of QSAR as a quantitative means for formulating enzyme mutations designed to achieve enhanced HRP reactivity towards compounds of specific interest.

Biotechnol. Bioeng. 2007;98: 295–299.

© 2007 Wiley Periodicals, Inc.

KEYWORDS: QSAR; site-directed mutagenesis; HRP

Introduction

We previously reported formulation of a robust quantitative structure–activity relationship (QSAR) for predicting initial rates at which aqueous-phase phenolic chemicals are degraded by the enzyme horseradish peroxidase (HRP) (Colosi et al., 2006). Normalized reaction rate constants (k_{CAT}) were found to correlate quantitatively with two independent variables. The first, the energy of the substrate's highest occupied molecular orbital (E_{HOMO}), was selected at the outset of the modeling effort because HRP-mediated

reactions involve electron transfer from substrate to the heme and this parameter defines the intrinsic redox reactivity of a substrate. As anticipated, E_{HOMO} was found to correlate strongly with k_{CAT} for the majority of fifteen different chemical substrates investigated. Data for these substrates are represented in Figure 1 using darkened squares (■). The remaining substrates, those deviating from the $E_{\text{HOMO}} - \ln(k_{\text{CAT}})$ trend, are represented by the open squares (□) and are labeled with an abbreviation for the compound to which they correspond. As a group, the latter compounds exhibit a negative residual and comprise larger, “bulkier” substrate molecules, indicating the need of additional QSAR parameters for better modeling. Figure 1 suggests, while the rate at which a phenolic compound may be degraded by HRP is intrinsically determined by its E_{HOMO} value, reaction rate for a compound too large or sterically hindered to facilitate optimal binding with the enzyme will be decreased relative to its maximum potential value.

Computational simulation, in which a model of the HRP active site was used to investigate binding of a phenolic substrate, was undertaken to characterize the manner in which molecular bulk mediates a decrease in initial reaction rate with HRP. Specifically, an energy-minimization algorithm was used to evaluate two intermolecular distances within the enzyme–substrate complex. The first distance, between the phenolic proton and the heme iron, exhibited no correlation with measured reaction rates in this study or two that preceded it (DeRiso et al., 2003; Van Haandel et al.,

Correspondence to: W.J. Weber Jr.

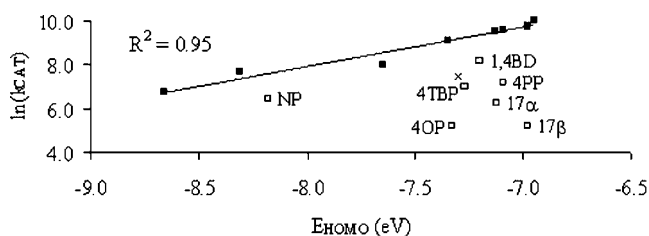


Figure 1. Degradation rate constants as a function of E_{HOMO} for the 15 phenolic chemicals included in the original analysis plus one additional chemical, 2-methoxyphenol (\times). For ease of reference, compounds deviating from the $E_{\text{HOMO}} - \ln(k_{\text{CAT}})$ trend are labeled with an abbreviation for the chemical to which they correspond: 4-octylphenol (4OP), 17 β -estradiol (17 β), 17 α -ethynylestradiol (17 α), 4-nitrophenol (4NP), 1,4-benzenediol (1,4BD), 4-*tert*-butylphenol (4TBP), and 4-phenylphenol (4PP). Error bars represent ± 1 SE of regression. (Figure is adapted from Colosi et al. (2006).)

1996). In contrast, the energy-weighted average distance between the phenolic proton and the purported binding site (δN on HIS42's imidazole ring) (Poulos and Kraut, 1980) was found to be well-correlated with deviation from the $E_{\text{HOMO}} - \ln(k_{\text{CAT}})$ trend. Thus, the use of the average characteristic binding distance (μ_{HIS}) as a second QSAR parameter unifies the two classes of chemical substrates investigated as indicated in Figure 2 (Colosi et al., 2006). This is consistent with the current understanding of the “push–pull” mechanism by which HRP Compound I is thought to oxidize a phenolic substrate, whereby formation of a hydrogen bond and subsequent proton transfer between the substrate proton and the δN of the HIS42 imidazole facilitates electron transfer to the heme (Al-Mustafa and Kincaid, 1994; Dunford, 1991).

Our ultimate goal is to use the QSAR as a design tool to achieve mutant HRP strains with enhanced degradation capacity compared to the wild type. Here we report an important first step towards that goal; i.e., validation of the hypothesis that reduction in calculated binding distances between substrate and a mutant HRP strain is associated with a proportional increase in enzyme reactivity.

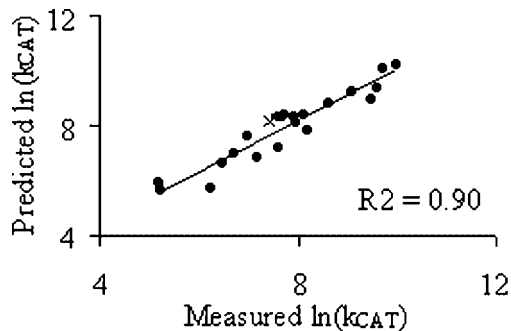


Figure 2. Verification that all compounds are unified using the QSAR model. Because 2-methoxyphenol (\times) also fits the unified trend, it is well-suited for validation of the model as a design tool.

Materials and Methods

Materials

Extracellular horseradish peroxidase (type I, RZ = 1.3), hydrogen peroxide (29.9%, ACS reagent grade), and 2-methoxyphenol (guaiacol) were purchased from Sigma Chemical Co. (St. Louis, MO).

Measurement of 2-Methoxyphenol Reaction Rate

A method reported earlier was adopted to measure the initial reaction rate of 2-methoxyphenol (Colosi et al., 2006). In brief, reactions mediated by HRP/ H_2O_2 were carried out at room temperature in 5 mL of phosphate buffer (10 mM, pH 7.0) using 7 mL Teflon-capped glass test tubes shaken vigorously over the reaction period to ensure their behavior as completely mixed batch reactors (CMBRs), and thus to allow the rate relationships described below for data interpretation (Weber and DiGiano, 1996). Various initial concentrations (S_0) of 2-methoxyphenol substrate were used at a fixed HRP dosage of 0.1 unit/mL, where one unit of HRP activity is defined as the quantity catalyzing the oxidation of 1 μmol of ABTS per minute. The reaction in each CMBR was initiated by addition of 150 μM H_2O_2 , an amount determined in preliminary tests as sufficient to saturate the enzyme. Each CMBR was shaken by hand for 20 s prior to the addition of ~ 100 μL of 1.0 N HCL to terminate the reaction. Three replicate experiments and a blank (equivalent volume buffer used in place of peroxide) were performed for each initial substrate concentration investigated. Following termination of the reaction and centrifugation at 2205g for 25 min, supernatant from each tube was transferred to an amber HPLC vial for determination of pre-reaction (blank = S_0) and post-reaction (S_t) concentrations of parent substrate. These determinations were made using an Agilent 1100 series high-performance liquid chromatograph (HPLC) equipped with a Phenomenex C18 column (250 mm \times 2.0 mm, 5 μm particle size). The mobile phase consisted of 40% reagent-grade acetonitrile (ACN) and 60% distilled deionized water (DDI) flowing at 0.75 mL/min, and the retention time of 2-methoxyphenol was 5.9 min.

Following measurement of pre-reaction (S_0) and post-reaction substrate (S_t) concentrations, the initial reaction rate (r_0) was defined using the formulation $r_0 = (S_0 - S_t)/\Delta t$, where Δt was always 20 s. Calculated values for r_0 were plotted as a function of S_0 , and the data were fit to the classical Michaelis–Menten equation according to $r_0 = r_{\text{MAX}} S_0/(K_M + S_0)$, where r_{MAX} is the maximum rate of reaction and K_M is the substrate's Michaelis constant. By definition, the maximum reaction rate, $r_{\text{MAX}} = k_{\text{CAT}}[E]$, is achieved when substrate concentration is sufficiently large to saturate all available enzyme molecules ($S_0 \gg K_M$). Although the technical reality is that $[E]$ is reduced over time as the enzyme becomes inactivated, we assumed that the 20 s reaction period used to determine the initial reaction rate

was sufficiently short such that $[E]$ could always be approximated by $[E_0]$. We thus calculated a normalized initial reaction rate constant according to $k_{\text{CAT}} = r_{\text{MAX}}/[E_0]$.

Computational Simulation of Interactions between 2-Methoxyphenol and HRP

Computations were performed as described in detail by Colosi et al. (2006) using the molecular and quantum mechanics algorithms available as part of HyperChem Molecular Modeling System, release 7.1, and are here outlined only in brief. Preliminary geometry optimization for 2-methoxyphenol was achieved using the OPLS molecular mechanics force field and the Polake-Ribiere optimization. Computations were performed in vacuo at 298 K, even though empirical measurements of initial reaction rate were measured in phosphate-buffered saline at room temperature. The in vacuo computation is regarded as an appropriate approach to approximate the active site, which, in this case, comprised only the interior portion of the HRP molecule shielded from interaction with solvent molecules by the rest of the enzyme. Further, the OPLS force field has been previously demonstrated effective for use with enzymes containing calcium and iron atoms, including HRP, with satisfactory results (Zelent et al., 2004). Subsequent quantum optimization for electronic structure was achieved using the ZINDO/1 semiempirical method, also with the Polake-Ribiere optimization. Selected molecular parameters were determined using the resulting optimized substrate structure.

The full structural coordinates of a model horseradish peroxidase, Compound I (C1A, entry 1HCH), were downloaded from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) (Berglund et al., 2002; Berman et al., 2001). Amino acid residues having no documented role in substrate binding were then eliminated (Howes et al., 2001; Thanabal et al., 1987; Xie et al., 2002), but crystallographic water molecules and two calcium ions were retained. Appropriate charge distribution for HRP Compound I was computed for the subset structure comprising the porphyrin ring, the heme iron, its associated oxygen, and proximal residue HIS170 using ZINDO/1 unrestricted Hartree–Fock (UHF) optimization (Wirstam et al., 1999). In simulations involving mutant strains of

HRP, pertinent amino acids were changed from the PDB downloaded structure.

To simulate interaction, the computational model for the enzyme was merged with that for 2-methoxyphenol, and the substrate aligned in a random location within the enzyme's distal region. The OPLS molecular mechanics force field was then utilized in a 1,000-step Monte Carlo (MC) simulation. All model components except the substrate molecule and two distal residues thought to be most critical for electron transfer (HIS42) and charge stabilization (ARG38) were held fixed during simulated substrate docking (Rodriguez-Lopez et al., 1996, 2001), and three pieces of information were recorded for each step of the simulation: the potential energy corresponding to the ensemble conformation, the distance between the substrate proton and the heme iron, and the distance between the substrate proton and δN on the HIS42 imidazole. Following completion of a 1,000-step MC simulation, the lowest energy conformation was recalled and its substrate geometry optimized using OPLS and Polake-Ribiere. The resultant optimized geometry was then used as the starting configuration for the next 1,000-step MC simulation. This process, MC simulations alternating with OPLS optimizations of the substrate molecule alone, was iterated until the energy decrease between two consecutive geometry-optimized structures was less than 1 kcal/mol. Additional starting configurations, followed by the alternating MC/OPLS iterations, were chosen at random until 50,000 steps had been sampled.

A range of binding distances was thus collected over the course of each docking simulation. To determine the characteristic binding distance of a particular interaction, the ranges were synthesized into an energy-weighted average distance, where the weights were assigned according to the potential energy value associated with each enzyme/substrate conformation according to the exponential distribution. Use of the energy weighting factors allowed lower-energy conformations, which are more favorable and thus more probable, to be weighted more heavily in the computation of the average distance compared to less favorable, higher-energy conformations.

Results and Discussion

In our earlier work (Colosi et al., 2006) we demonstrated that the reactivity of HRP with a phenolic chemical is a

Table 1. Normalized reactivity and computed binding distance for reaction between 2-methoxyphenol and seven strains of HRP Compound I including the wild type and six mutants.

HRP strain	Data source	$\ln(k_{\text{CAT}})$ (s^{-1})	μ_{HIS} (\AA)
WT	This study	7.42	7.46
L131P	Morawski et al. (2001)	7.73	7.23
L131P + L223Q	Morawski et al. (2001)	8.64	6.73
L131P + A175S	Morawski et al. (2001)	7.59	7.26
F41L	Ozaki and Ortiz de Montellano (1995)	8.12	7.20
F41T	Ozaki and Ortiz de Montellano (1995)	7.91	7.29
F41A	Newmyer and Ortiz de Montellano (1995)	7.70	7.30

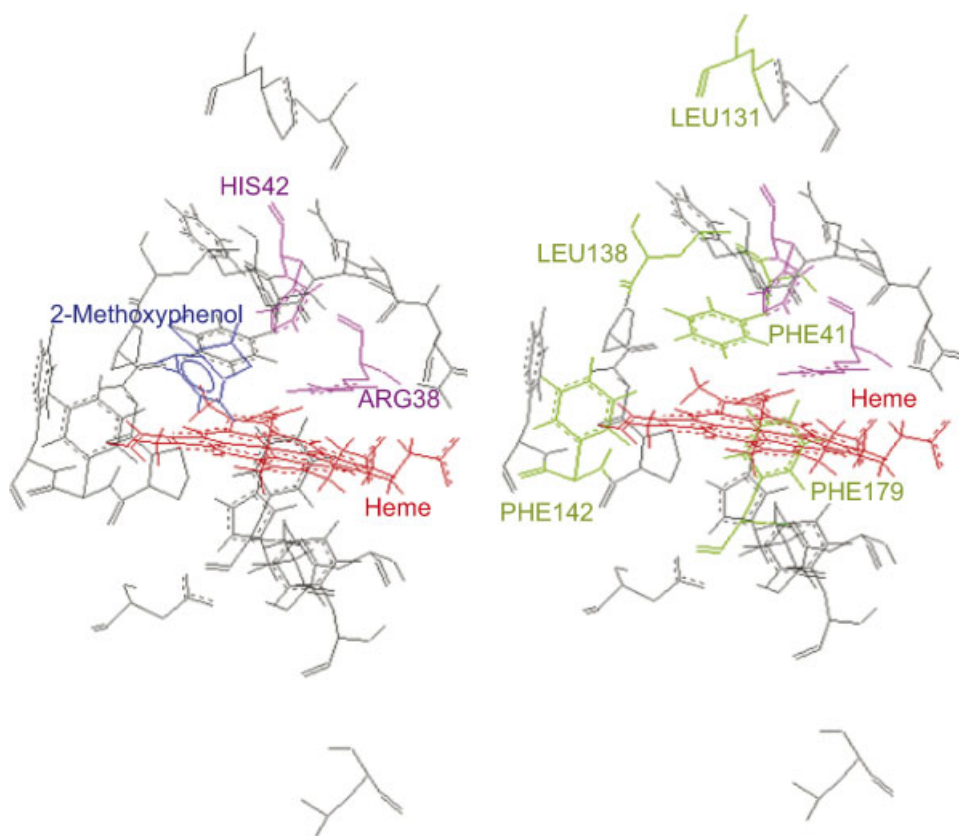


Figure 3. A representative docking ensemble, 2-methoxyphenol + HRP. On the left, relative locations of 2-methoxyphenol, the heme, and two distal residues thought to be most critical in substrate docking (ARG38 and HIS42). On the right, locations of mutated amino acid residues relative to the heme in the same viewing orientation. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

function of two independent parameters; i.e., E_{HOMO} and the characteristic binding distance μ_{HIS} . The value of the former is fixed for a given substrate, thus enhanced degradation of a target compound is theoretically achievable only if its characteristic binding distance can somehow be

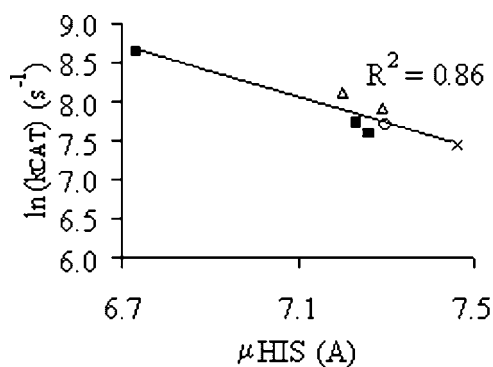


Figure 4. Measured values of $\ln(k_{\text{CAT}})$ as a function of simulated μ_{HIS} distance for wild type and six mutants strains of HRP Compound I. Individual data sources are indicated with different symbols: wild type reference (x), Morawski et al. (2001) (■), Ozaki and Ortiz de Montellano (1995) (▲), and Newmyer and Ortiz de Montellano (1995) (○).

reduced. In this communication we demonstrate the possibilities for use of QSAR modeling as a design tool by confirming the validity of our hypothesis that changes in HRP primary structure are in fact able to mediate both a reduction in its characteristic binding distance for a given substrate and its associated increase in k_{CAT} in a manner predicted by the two-parameter QSAR model presented earlier.

To verify the above-mentioned hypothesis, we utilize previously published data for degradation of 2-methoxyphenol by HRP mutants produced and evaluated in other laboratories. The substrate 2-methoxyphenol, which is frequently used as an indicator in a colorimetric assay for HRP activity, was not one of the fifteen chemicals included in our original QSAR analysis. We specifically selected it for use in the investigation reported here because measurements of its degradation rate by HRP mutants are widely available from other sources. We then measured its wild-type HRP degradation rate ($\ln(k_{\text{CAT}}) = 7.42 \text{ s}^{-1}$) and computed the associated wild-type HRP binding distance ($\mu_{\text{HIS}} = 7.46 \text{ \AA}$) using the same protocols employed in formulating the original QSAR. As indicated in Figure 2, the reactivity of 2-methoxyphenol can be predicted well using the two-parameter QSAR model we had developed earlier. Moreover, its reactivity is, as shown in Figure 1, less

than what would be predicted based on the $E_{\text{HOMO}} - \ln(k_{\text{CAT}})$ trend, indicating that its binding with the wild-type HRP is sub-optimal. This confirms that 2-methoxyphenol is indeed a good model substrate for use in evaluating the hypothesis that HRP mutations that reduce μ_{HIS} can improve degradation efficiency.

Literature measurements of mutant $\ln(k_{\text{CAT}})$ values (Morawski et al., 2001; Newmyer and Ortiz de Montellano, 1995; Ozaki and Ortiz de Montellano, 1995), relative to their respective recombinant wild types, were available for six HRP strains: L131P (a proline replaces leucine 131), L131P + L223Q (a proline replaces leucine 131 and a glutamine replaces leucine 223), L131P + A175S (a proline replaces leucine 131 and a serine replaces alanine 175), and F41A, F41L, and F41T (in which phenylalanine 41 is replaced by an alanine, leucine, and threonine, respectively). Figure 3 shows the location of these amino acids relative to the enzyme's active site. To compensate for differences in experimental methodologies, units, etc. among the three different data sources, measurements of 2-methoxyphenol degradation rate made by other authors' mutants were normalized using the ratio of their wild type rate measurement to our own. Table I outlines the resulting normalized degradation rates, which are presented alongside the characteristic binding distances computed during simulated interaction with 2-methoxyphenol. Figure 4 depicts this data set graphically.

As illustrated clearly in Figure 4, our findings indicate that the rate of HRP-mediated 2-methoxyphenol degradation is directly proportional to the binding distance, just as predicted by our earlier two-parameter QSAR model. This figure thus validates the use of μ_{HIS} as a legitimate design parameter; i.e., genetic alterations of HRP's primary structure that mediate decreases in binding distances with respect to a particular substrate are indeed associated with increased specific activity with that substrate. Finally, we see that changes in reactivity effected by changes in binding distance occur in a quantitative, linear manner, consistent with expectations based on the $E_{\text{HOMO}} - \mu_{\text{HIS}}$ QSAR. The slope of the trace presented in Figure 4 suggests that each one-angstrom decrease in enzyme-substrate binding distance mediates a 1.6 unit increase in the $\ln(k_{\text{CAT}})$ value. As we might expect, this reflects sensitivity on the same order of magnitude as predicted by the slope of the previously developed two-parameter QSAR model; i.e.,: $\ln(k_{\text{CAT}}) = 25.4 + 1.4E_{\text{HOMO}} - 0.9\mu_{\text{HIS}}$.

The results comprising Figure 4 validate the intuitive notion that not all active site mutations will affect binding distances, and, by extension enzyme reactivity, in exactly the same manner. This has significant implications for QSAR-assisted rational protein re-design, a potentially powerful tool for targeted degradation of selected chemical species. Ultimately, it will be necessary to design mutations that effect a reduction in mutant binding distance without compromising protein functionality in order to achieve enhanced HRP reactivity towards compounds of interest using strategies of site-directed mutagenesis.

This research was financed in major part by Research Grant P42ES04911-14 from the National Institutes for Environmental and Health Sciences. LMC much appreciates additional support provided to her directly by a National Science Foundation Graduate Research Fellowship.

References

- Al-Mustafa J, Kincaid JR. 1994. Resonance Raman study of cyanide-ligated horseradish peroxidase. Detection of two binding geometries and direct evidence for the "push-pull" effect. *Biochemistry* 33:2191-2197.
- Berglund GI, Carlsson GH, Smith AT, Szoke H, Henriksen A, Hajdu J. 2002. The catalytic pathway of horseradish peroxidase at high resolution. *Nature* 417:463-468.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2001. The protein data bank. *Nucleic Acids Res* 28:235-242.
- Colosi LM, Huang Q, Weber WJr. 2006. Quantitative structure-activity relationship based quantification of the impacts of enzyme-substrate binding on rates of peroxidase-mediated reactions of estrogenic phenolic chemicals. *J Am Chem Soc* 128:4041-4047.
- DeRiso A, Gullotti M, Casella L, Monzani E, Profumo A, Gianelli L, DeGioia L, Gaiji N, Colonna S. 2003. Selectivity in the peroxidase catalyzed oxidation of phenolic sulfides. *J Mol Catal A* 204:391-400.
- Dunford HB. 1991. *Peroxidases in chemistry and biology*, Vol. 2. Boca Raton, FL: CRC Press.
- Howes BD, Veitch NC, Smith AT, White CG, Smulevich G. 2001. Haem-linked interactions in horseradish peroxidase revealed by spectroscopic analysis of the Phe 221 → Met mutant. *Biochem J* 353:181-191.
- Morawski B, Quan S, Arnold FH. 2001. Functional expression and stabilization of horseradish peroxidase by directed evolution in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 76:99-107.
- Newmyer SL, Ortiz de Montellano PR. 1995. Horseradish peroxidase his-42 → ala, his-42 → val, and phe-41 → ala mutants. *J Biol Chem* 270:19430-19438.
- Ozaki S, Ortiz de Montellano PR. 1995. Molecular engineering of horseradish peroxidase: Thioether sulfoxidation and styrene epoxidation by phe-41 leucine and threonine mutants. *J Am Chem Soc* 117:7056-7064.
- Poulos TL, Kraut J. 1980. The stereochemistry of peroxidase catalysis. *J Biol Chem* 255:8199-8205.
- Rodriguez-Lopez JN, Smith AT, Thorneley RNF. 1996. Role of arginine in horseradish peroxidase. *J Biol Chem* 271:4023-4030.
- Rodriguez-Lopez JN, Lowe DJ, Hernandez-Ruiz J, Hiner ANP, Garcia-Canovas F, Thorneley RNF. 2001. Mechanism of reaction of hydrogen peroxide with horseradish peroxidase: Identification of intermediates in the catalytic cycle. *J Am Chem Soc* 123:11838-11847.
- Thanabal V, de Ropp JS, LaMar GN. 1987. ¹H NMR study of the electronic and molecular structure of the heme cavity in horseradish peroxidase. Complete heme resonance assignments based on saturation transfer and nuclear overhauser effects. *J Am Chem Soc* 109:265-272.
- Van Haandel MJH, Rietjens IMCM, Soffers AEMF, Veeger C, Vervoort J, Modi S, Mondal MS, Patel PK, Behere DV. 1996. Computer calculation-based quantitative structure-activity relationships for the oxidation of phenol derivatives by horseradish peroxidase compound II. *J Biol Inorg Chem* 1:460-467.
- Weber WJr, DiGiano FA. 1996. *Process dynamics in environmental systems*. New York: John Wiley & Sons. 600 p.
- Wirstam M, Blomberg MRA, Siegbahn PEM. 1999. Reaction mechanism of compound I formation in heme peroxidases: A density functional theory study. *J Am Chem Soc* 121:10178-10185.
- Xie Y, Das PK, Caaveiro JMM, Klibanov AM. 2002. Unexpectedly enhanced stereoselectivity of peroxidase-catalyzed sulfoxidation in branched alcohols. *Biotechnol Bioeng* 79:105-111.
- Zelent B, Kaposi AD, Nucci NV, Sharp KA, Dalosto SD, Wright WW, Vanderkooi JM. 2004. Water channel of horseradish peroxidase studied by the charge-transfer absorption band of ferric heme. *J Phys Chem* 108:10317-10324.