J|A|C|S
ARTICLES
Published on Web 03/08/2006

Quantitative Structure−Activity Relationship Based Quantification of the Impacts of Enzyme−Substrate Binding on Rates of Peroxidase-Mediated Reactions of Estrogenic Phenolic Chemicals

Lisa M. Colosi,[†] Qingguo Huang,[‡] and Walter J. Weber, Jr.*[,‡]

Contribution from the Environmental and Water Resources Engineering Program and Department of Chemical Engineering, University of Michigan, Ann Arbor, Michigan 48109-2099

Received October 31, 2005; E-mail: wjwjr@umich.edu

Abstract: The initial rates of horseradish peroxidase (HRP)-mediated enzymatic reactions of 15 assorted aqueous phase phenolic chemicals were studied. The associated reaction rate constants were found to correlate quantitatively with two independent variables: the highest-occupied molecular orbital energy (E_{HOMO}) defining the intrinsic redox reactivities of the phenolic substrates and the distance between a substrate and the δN of HIS42's imidazole ring in an HRP/substrate binding complex, obtained through molecular simulations. Highly correlated quantitative structure−activity relationship (QSAR) equations were thus developed. This work provides insights into the impacts that HRP/substrate binding may have on HRP-mediated reactions. Additionally, the QSAR equations developed in the work may serve as a basis to further explore the potential use of HRP-mediated reactions in the treatment of estrogenic contaminants, and they constitute an important tool for redesign and screening of potential proteomic modifications to the wild-type HRP structure intended to enhance reactivity toward selected substrates.

Introduction

As an "emerging" class of contaminants, endocrine disrupting chemicals (EDCs) have become important foci of public concern and scientific interest. Such chemicals mimic and interfere with hormonal activities and thus adversely impact ecosystems and human health by disrupting reproduction, growth, and development.[1,2] Studies have also linked chronic ingestion of EDCs to cancer.[3−5] Many EDCs tend to be persistent and bioaccumulative in the environment and are capable of eliciting physiological response even at extremely low concentrations (0.1−10 ng/L).[6,7] This combination of chemical characteristics, which allows them to mediate cellular activities, also makes them a challenge from the perspective of water/wastewater treatment.[8,9] It has been demonstrated that many EDCs are either resistant to or only partially degraded in traditional water and wastewater treatment processes and thus may persist at potent levels in treated effluents.[7,8,10]

One particularly important class of EDCs comprises those chemicals capable of binding to natural estrogen receptors (e.g., ER-α) and eliciting and/or interfering with estrogenic responses normally associated with endogenous estrogens (e.g., 17β-estradiol).[4] These chemicals do not necessarily share a close structural resemblance to natural estrogens.[11] Instead, binding to the estrogen receptors is thought to be relatively nonspecific so long as the substrate possesses a phenolic substructure, with a hydrophobic moiety at the para position.[12] Thus, certain substituted phenols and their derivatives comprise an important class of estrogenic contaminants.[13]

It is intriguing that many phenolic chemicals are also active substrates of the enzyme horseradish peroxidase (HRP, EC 1.11.1.7; donor H_2O_2). This enzyme mediates the reaction of phenolic substrates in a three-step catalytic cycle (Figure 1). As depicted in Figure 1, the ground-state HRP is oxidized by H_2O_2 in the first step to yield an enzyme intermediate (compound I), which in the second step abstracts an electron from the phenolic substrate AH_2 to become another enzyme intermediate (compound II) and produce a phenoxy radical $AH^{\bullet}$.

[†] Environmental and Water Resources Engineering Program.
[‡] Department of Chemical Engineering.

(1) Purdom, C. E.; Hardiman, P. A.; Bye, V. J.; Eno, N. C.; Tyler, C. R.; Sumpter, J. P. *Chem. Ecol.* **1994**, *8*, 275.
(2) Kavlock, R. J., et al. *Environ. Health Perspect.* **1996**, *104*, 715. (A full list of authors is given in the Supporting Information IV.)
(3) Jobling, S.; Nolan, M.; Tyler, C. R.; Brighty, G.; Sumpter, J. P. *Environ. Sci. Technol.* **1998**, *32*, 2498.
(4) Saliner, A. G.; Amat, L.; Carbo-Dorca, R.; Schultz, T. W.; Cronin, M. T. D. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1166.
(5) Choi, S. M.; Yoo, S. D.; Lee, B. M. *J. Toxicol. Environ. Health, Part B* **2004**, *7*, 1.
(6) Welshons, W. V.; Thayer, K. A.; Judy, B. M.; Taylor, J. A.; Curran, E. M.; vom Saal, F. S. *Environ. Health Perspect.* **2003**, *111*, 994.
(7) Baronti, C.; Curini, R.; D'Ascenzo, G.; DiCorcia, A.; Gentili, A.; Samperi, R. *Environ. Sci. Technol.* **2000**, *34*, 5059.
(8) Johnson, A. C.; Sumpter, J. P. *Environ. Sci. Technol.* **2001**, *35*, 4697.
(9) Snyder, S. A.; Westerhoff, P.; Yoon, Y.; Sedlak, D. L. *Environ. Eng. Sci.* **2003**, *20*, 449.
(10) Desbrow, C.; Routledge, E. J.; Brighty, G. C.; Sumpter, J. P.; Waldock, M. *Environ. Sci. Technol.* **1998**, *32*, 1549.
(11) Petrovic, M.; Eljarrat, E.; de Alda, M. J. L.; Barcelo, D. *TrAC Trends Anal. Chem.* **2001**, *20*, 637.
(12) Fang, H.; Tong, W.; Shi, L. M.; Blair, R.; Perkins, R.; Branham, W.; Hass, B. S.; Xie, Q.; Dial, S. L.; Moland, C. L.; Sheehan, D. M. *Chem. Res. Toxicol.* **2001**, *14*, 280.
(13) Petrovic, M.; Eljarrat, E.; de Alda, M. J. L.; Barcelo, D. *J. Chromatogr., A* **2002**, *974*, 23.

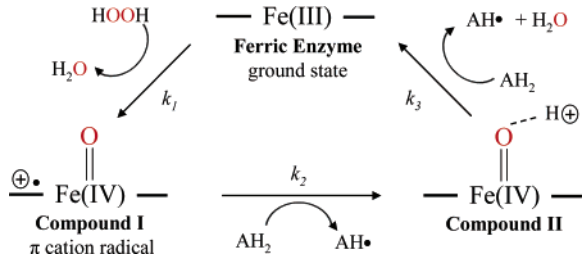

Figure 1. Schematic depiction of the three-step HRP catalytic cycle (adapted from Berglund et al.[16]).

In the third step, compound II abstracts another electron from yet another AH_2 to produce a second phenoxy radical, at which point the enzyme converts back to the ground state. It is known that the first and second reactions occur much more rapidly than the third (k_1, $k_2 \gg k_3$) and that the third reaction is thus rate limiting.[14−16]The phenoxy radicals generated during these HRP-mediated reactions couple with each other to form polymeric products that can be readily removed from water or soil.[17−20] In consideration of the generally high specificity and efficiency of enzymatic reactions, such HRP-mediated reactions constitute a potentially attractive means for treatment of waters containing estrogenic phenolic chemicals at low levels.[21]

In this study, we tested the initial reaction rates of 15 assorted phenols mediated by HRP in the aqueous phase. The overall reaction rate at substrate saturation conditions, rather than rates for each intermediate reaction step, was measured to reflect the enzyme activity specific to each chemical that could be anticipated in water treatment applications. The results indicate that the enzyme catalytic rate constant (k_{cat}) and specificity (K_M) vary markedly for the phenolic substrates studied. This prompted us to explore quantitative structure−activity relationships (QSAR) that effect such significant variability in reactivity.

Earliest QSAR studies for HRP-mediated reactions have concentrated on correlations between substrate molecular structures and their reaction rates, either for individual intermediate steps or for the overall reaction cycle. Because of the nature of the redox reaction, four molecular descriptors have most frequently been used: i.e., energy of the highest-occupied molecular orbital (E_{HOMO}), energy of the lowest-unoccupied molecular orbital (E_{LUMO}), atomic charge distribution, and the Hammet constant (σ). Improved computation power has made *ab initio* and semiempirical estimates of E_{HOMO} increasingly accurate, and this descriptor has been correlated, to varying extent of fit, with reaction rates in numerous investigations for various series of substrates. Reported correlation coefficients, r, are on the range 0.560−0.998.[15,22−28] The reported extent of correlation with E_{LUMO} has been inconsistent. Sakurada and Brewster reported fairly strong correlations (r = 0.86, 0.89) between compound II reactivity and E_{LUMO} for sets of substituted phenols, whereas Hosoya reported no significant correlation for a set of similar substrates.[23,24,28] Despite early work by Bordeleau and Bartha indicating a correlation between compound II reactivity and atomic charge on the substrate's phenolic oxygen, Hosoya and Sakurada were unable to confirm significant correlation.[23,28,29] Job, Dunford, Sakurada, and Gilabert have reported significant correlations between the Hammet constant σ of the substrates and their reaction rates with both compounds I and II.[15,23,27,30]

The studies described above, taking a traditional QSAR approach, focused on correlating reactivities to molecular or electronic structures of the substrates but neglected enzyme−substrate binding interactions. An implicit assumption involved with this approach is that all substrates under study have a similar conformation of transition state (enzyme−substrate complex) so that the activation energy linearly correlates with the substrate states. This assumption, apparently, does not easily hold true for enzymatic reactions in which protein molecules are both complex and flexible, and as such, QSARs for enzymatic reactions are often valid only within a limited set of substrates with only small structural variations. Incorporation of key descriptors of enzyme−substrate binding features may make the QSAR a more powerful tool to investigate enzymatic reactions. Strategic binding distances can often serve as a convenient quantitative parametrization of the spatial conformation of enzyme−substrate complexes.[25,31] van Haandel and de Riso have both tried to correlate HRP-mediated reaction rates to NMR-measured distances between the substrate proton and the heme iron.

HRP, like other heme peroxidases, contains at its catalytic center an iron porphyrin with reactivity modulated by surrounding amino acids comprising the catalytic pocket.[16] One widely accepted reaction mechanism put forth by Poulos and Kraut, corroborated by measurements using nuclear magnetic resonance and rapid-scan spectrophotometry, suggests that electron transfer from the substrate to the heme takes place not at the iron atom itself but at the δN of HIS42's imidazole ring.[32] Distal residue ARG38 works in tandem with HIS42, providing crucial charge stabilization during electron transfer and facilitating both the oxidation and the simultaneous delocalization of the electron over the entire porphyrin ring.[33,34]

In this study, we found that the markedly different reactivities of the 15 phenols, many of them documented or suspected estrogens, cannot be singularly correlated to the substrate

(14) Dunford, H. B. Horseradish peroxidase: structure and kinetic properties. In *Peroxidases in Chemistry and Biology*; Everse, J. E., Everse, K. E., Grisham, M. B., Eds.; CRC Press: Ann Arbor, MI, 1991; Vol. II, pp 1−24.
(15) Job, D.; Dunford, H. B. *Eur. J. Biochem.* **1976**, *66*, 607.
(16) Berglund, G. I.; Carlsson, G. H.; Smith, A. T.; Szoke, H.; Henriksen, A.; Hajdu, J. *Nature* **2002**, *417*, 463.
(17) Shiga, T.; Imaizumi, K. *Arch. Biochem. Biophys.* **1975**, *167*, 469.
(18) Klibanov, A. M.; Tu, T. M.; Scott, K. P. *Science* **1983**, *221*, 259.
(19) Caza, N.; Bewtra, J. K.; Biswas, N.; Taylor, K. E. *Water Res.* **1999**, *33*, 3012.
(20) Dec, J.; Bollag, J. M. *J. Environ. Qual.* **2000**, 29, 665.
(21) Huang, Q.; Weber, W. J., Jr. *Environ. Sci. Technol.* **2005**, *36*, 6029.
(22) Sakurada, J.; Alda, M.; Nagata, C.; Hosoya, T. *J. Biol. Phys.* **1988**, *16*, 17.
(23) Sakurada, J.; Sekiguchi, R.; Sato, K.; Hosoya, T. *Biochemistry* **1990**, *29*, 4093.
(24) Brewster, M. E.; Doerge, D. R.; Huang, M. J.; Kaminski, J. J.; Pop, E.; Bodor, N. *Tetrahedron* **1991**, *47*, 7525.
(25) Van Haandel, M. J. H.; Rietjens, I. M. C. M.; Soffers, A. E. M. F.; Veeger, C.; Vervoort, J.; Modi, S.; Mondal, M. S.; Patel, P. K.; Behere, D. V. *J. Biol. Inorg. Chem.* **1996**, *1*, 460.
(26) Van Haandel, M. J. H.; Claassens, M. M. J.; Van der Hout, N.; Boersma, M. G.; Vervoort, J.; Rietjens, I. M. C. M. *Biochim. Biophys. Acta* **1999**, *1435*, 22.
(27) Gilabert, M. A.; Hiner, A. N. P.; Garcia-Ruiz, P. A.; Tudela, J.; Garcia-Molina, F.; Acosta, M.; Garcia-Canovas, F.; Rodriguez-Lopez, J. N. *Biochim. Biophys. Acta* **2004**, *1699*, 235.
(28) Hosoya, T.; Fuji, T.; Ogawa, S. *J. Theor. Biol.* **1983**, *100*, 283.
(29) Bordeleau, L. M.; Bartha, R. *Can. J. Microbiol.* **1972**, *18*, 1873.
(30) Dunford, H. B.; Adeniran, A. J. *Arch. Biochem. Biophys.* **1986**, *251*, 536.
(31) De Riso, A.; Gullotti, M.; Casella, L.; Monzani, E.; Profumo, A.; Gianelli, L.; De Gioia, L.; Gaiji, N.; Colonna, S. *J. Mol. Catal. A* **2003**, *204*, 391.
(32) Poulos, T. L.; Kraut, J. *J. Biol. Chem.* **1980**, *255*, 8199.
(33) Rodriguez-Lopez, J. N.; Smith, A. T.; Thorneley, R. N. F. *J. Biol. Chem.* **1996**, *271*, 4023.
(34) Rodriguez-Lopez, J. N.; Lowe, D. J.; Hernandez-Ruiz, J.; Hiner, A. N. P.; Garcia-Canovas, F.; Thorneley, R. N. F. *J. Am. Chem. Soc.* **2001**, *123*, 11838.

structural variations. We computationally investigated the enzyme−substrate binding conformations and found that the distance between the substrate phenolic proton and δN of HIS42's imidazole ring plays an important role in modulating the substrate reactivities. By incorporation of such a parameter for binding conformation, QSAR equations were yielded to well describe the variation in reactivity of different substrates tested. This QSAR study provides insights into how enzyme/substrate binding interaction impacts HRP reactivity. As such, the results serve as a basis for design and screening of potential proteomic modifications to HRP for enhanced reactivity toward selected contaminants of concern, provided they are structurally similar to the substituted phenols evaluated in this study.

Experimental Section

Materials. Extracellular horseradish peroxidase (type I, RZ = 1.3), hydrogen peroxide (29.9%, ACS reagent grade), 1,2,3-benzenetriol (97%), 1,4-benzenediol (99+%), and 2,2′-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) (98%, in diammonium salt form) were purchased from Sigma Chemical Co. (St. Louis, MO). Bisphenol A (99+%), 2,6-dimethoxyphenol (99%), 4-nitrophenol (98%), 4-*tert*-butylphenol (99%), 4-phenylphenol (97%), 4-octylphenol (99%), 17β-estradiol (97%), 17α-ethynylestradiol (98%), 4-methoxyphenol (99%), 4-ethoxyphenol (99%), and 4-ethylphenol (99%) were obtained from Aldrich Chemical Co. (Milwaukee, WI). Phenol (99%) and 4-chlorophenol (99+%) were obtained from Acros Chemical Company (New Jersey).

Measurements of Initial Reaction Rate. 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 as reactors, with various initial concentrations of a substrate and a consistent dosage of HRP varying between 0.01 and 0.5 unit/mL for different substrates. One unit of HRP activity is defined as the amount catalyzing the oxidation of 1 μmol of ABTS per minute. Reaction in each tube was initiated upon addition of 150 μM H_2O_2, which was determined to be sufficient to saturate the enzyme in preliminary tests. Each tube 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 were performed for each initial substrate concentration. In addition, a blank tube was created using an equivalent volume of buffer in place of the peroxide. Following completion of reaction and centrifugation at 2205*g* for 25 min, supernatant from each tube was transferred to amber HPLC vials for determination of prereaction (blank) and postreaction concentrations of the parent substrate.

An Agilent 1100 series high-performance liquid chromatography (HPLC) equipped with a Phenomenex C18 column (250 × 2.0 mm, 5-μm particle size) was used to detect each estrogenic substrate. All concentrations were measured using UV absorbance with external calibration. A mobile phase consisting of 40% reagent-grade acetonitrile (ACN) and 60% distilled deionized water (DDI) was used at a flow rate of 0.3 mL/min to detect 17α-ethynylestradiol (RT = 15.9 min), 17β-estradiol (RT = 8.7 min), bisphenol A (RT = 11.9 min), phenol (RT = 4.7 min), 4-phenylphenol (RT = 10.2 min), 4-*tert*-butylphenol (RT = 12.3 min), and 4-chlorophenol (RT = 7.6 min). Detection of 4-octylphenol (RT = 8.6 min) was accomplished using a mobile phase comprised of 80% ACN and 20% DDI at a flow rate equal to 0.3 mL/min. Detection of 4-ethylphenol (RT = 10.4 min), 4-methoxyphenol (RT = 3.9 min), 4-ethoxyphenol (RT = 5.7 min), 4-nitrophenol (RT = 3.6 min), and 2,6-dimethoxyphenol (RT = 7.8 min) was achieved using a mobile phase comprising 30% ACN and 70% DDI at a flow rate of 0.4 mL/min. Acetic acid was added at 1% v/v to a mobile phase comprising 20% ACN and 80% DDI for detection of 1,2,3-benzenetriol (RT = 12.8 min) at a flow rate equal to 0.3 mL/min. A similarly acidified mobile phase (1% v/v) comprising 5% ACN and 95% DDI was used to detect 1,4-benzenediol (RT = 4.2 min) at a flow rate of 0.3 mL/min.

Following measurement of the substrate concentration in both the blank (S_0) and reaction tubes (S_t), the initial reaction rate (r_0) was defined using the formulation $r_0 = (S_0 - S_t)/\Delta t$. The reaction time, Δt, was always 20 s because it was short enough to capture the pseudo-first-order degradation but still sufficiently long to ensure reliable handling and reproducible results. Calculated values for r_0 were plotted as a function of S_0 for each substrate, and each set of data fit to the classical Michaelis−Menten equation according to $r_0 = r_{MAX} \times S_0/(K_M + S_0)$, where r_{MAX} is the maximum rate of reaction and K_M is the substrate's Michaelis constant. The maximum reaction rate $r_{MAX} = k_{CAT}[E]$ was achieved when substrate concentration was sufficiently large to saturate all available enzymes ($S_0 \gg K_M$). Strictly speaking, $[E]$ may be reduced over time during the reaction as the enzyme becomes inactivated, but in electing to measure only the initial reaction rate, we assume the time duration is sufficiently short such that $[E]$ may always be approximated by $[E_0]$. We thus calculated the initial reaction rate constant according to $k_{CAT} = r_{MAX}/[E_0]$.

Computational Models for Substrates and HRP. All computations were performed using molecular and quantum mechanics algorithms available as part of HyperChem Molecular Modeling System, release 7.1 (Hypercube, Inc.: Gainesville, Florida). For substrate chemicals, preliminary geometry optimization for bond lengths, bond angles, and torsion angles was achieved using the OPLS molecular mechanics force field and the Polak−Ribiere optimization algorithm set to a convergence gradient criterion less than 0.1 kcal/(Å-mol). Subsequent quantum optimization for electronic structure was achieved using the ZINDO/1 semiempirical method and the Polake−Ribiere optimization set to the same convergence. Following determination of the optimal quantum structure, partial charge distribution, V_M (molecular volume), E_{HOMO} (energy of the highest-occupied molecular orbital), and E_{LUMO} (energy of the lowest-unoccupied molecular orbital) for each substrate were then computed.

The 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).[16,34] Apart from the porphyrin ring, the heme iron, and its double-bonded oxygen atom, all residues were removed except for those with a documented role in the binding mechanism: ALA71, ALA140, ARG38, ASN70, ASN72, ASP247, GLU64, GLY69, HIS42, HIS170, LEU138, PHE41, PHE68, PHE142, PHE143, PHE179, PHE221, PRO139, and PRO141.[36−38] Also included in the resultant catalytic pocket model were two calcium atoms located in the immediate vicinity of the pocket, and a single water molecule, HOH311, was included because of its association with the calcium atom proximal to the heme residue.[36−38]

Prior to using the compound I model in molecular simulations, we computationally reassigned an appropriate charge distribution for a subset structure comprising the porphyrin ring, the heme iron, its associated oxygen, and proximal residue HIS170. For this computation, we hydroxylated HIS170's terminal carbonyl group and protonated its terminal amino group where it had been previously bonded to residues not included in the subset model. The charge assignment was achieved using ZINDO/1 unrestricted Hartree−Fock (UHF) optimization on the subset components with the convergence criterion set to 0.01 kcal/(Å-mol). The total charge was set to zero to reflect the +1 charge on the

(35) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2001,** 28, 235−242.
(36) Howes, B. D.; Veitch, N. C.; Smith, A. T.; White, C. G.; Smulevich, G. *Biochem. J.* **2001**, *353*, 181.
(37) Xie, Y.; Das, P. K.; Caaveiro, J. M. M.; Klibanov, A. M. *Biotechnol. Bioeng.* **2002**, *79*, 105
(38) Thanabal, V.; de Ropp, J. S.; LaMar, G. N. *J. Am. Chem. Soc.* **1987**, *109*, 7516.

heme[39] and the −1 charge on HIS170 ($pK_a = 6.04$) as a result of deprotonation at pH = 7.00.[40] The multiplicity ($2S + 1$) was set to quartet to reflect the two unpaired electrons residing in the Fe═O group and one electron on the porphyrin.[39] Figure S1 of Supporting Information I depicts the resulting charge distribution for the selected subset of compound I. The calculated net charges on the heme iron (+0.10) and the heme-bound oxygen (−0.17) are in reasonable agreement with literature values for charges on the heme iron (+0.16) and the heme-bound oxygen (−0.19).[39] Charges for the remaining residues comprising the catalytic pocket model were left unchanged from their PDB downloaded values.

To achieve a suitable computational model of compound II, charge distribution was reassigned for the subset structure that comprises the same components that have been used in the charge assignment computation for compound I, in addition to HIS42. HIS42 was included because it is believed to form a hydrogen bond with the iron-bound oxygen in compound II. The ZINDO/1 UHF method was again used to assign charges on this subset structure. The total charge was set to zero to represent the −1 charge on HIS170's imidazole N, the zero charge on the heme, and the +1 charge of the proton bonded to HIS42's imidazole N and coordinated to the heme-bound oxygen. The multiplicity was set to triplet to reflect the two unpaired electrons of the Fe═O center.[41] Following the ZINDO/1 UHF optimization, the distance between the proton on HIS42's imidazole δN and the heme-bound oxygen was measured to be 7.1 Å, within the range specified by HyperChem as a hydrogen bond.[43] Figure S1 of Supporting Information I depicts the resulting charge distribution for the selected subset of compound II. Charges for the rest of the model were left unchanged from their PDB downloaded values.

Simulation of Enzyme−Substrate Interactions and Calculation of Average Binding Distances. To simulate interaction between a substrate and compound I, the computational model of the enzyme was merged with that of a substrate. The substrate was first aligned into a random location within the enzyme's distal region. The OPLS molecular mechanics force field was then utilized in a 1000-step Monte Carlo (MC) simulation during which all model components were held fixed in place except for the substrate molecule and two distal residues thought to be most critical during docking, HIS42 and ARG38. This and all subsequent Monte Carlo simulations were performed in vacuo at 300 K with a maximum trial atom movement distance equal to 0.05 Å. In each step, a randomly selected atom was moved a random distance in a random direction. Moves resulting in negative energy changes were selected. Moves resulting in positive energy changes were accepted with a probability defined by Boltzmann distribution, $f(x) = e^{-\Delta E/k_B T}$, where ΔE is the magnitude of the energy increase, k_B is the Boltzmann constant, and T is absolute temperature. For each step of the MC simulation, we recorded three pieces of information: the potential energy corresponding to each step's conformation, the distance between the substrate's phenolic proton and the heme-bound oxygen, and the distance between the substrate's phenolic proton and the imidazole δN on HIS42. For substrate molecules with more than one phenolic hydroxyl group, only the minimum distances between the substrate proton and the reference oxygen and nitrogen atoms were recorded. Following completion of the first 1000-step MC simulation, the lowest-energy conformation was recalled and its substrate geometry was optimized using OPLS and Polak−Ribiere. The resultant optimized geometry was then used as the starting configuration for the next 1000-step MC simulation in which the substrate molecule, HIS42, and ARG38 were once again allowed to move freely. 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. This generally required four to seven 1000-step MC simulations. Such an optimization process was repeated for the same substrate at seven to ten different starting positions that were randomly selected. Substrate/compound II binding was also optimized using the same approach for each of the 15 substrates.

A range of binding distances were collected for each substrate over the course of its docking simulation. To facilitate a QSAR analysis, the ranges were synthesized into an average distance weighted according to the potential energy value associated with each enzyme/substrate conformation according to the exponential distribution. The probability density function for the exponential distribution takes the form $f(x) = \lambda e^{-\lambda x}$, where x is a random variable equal to the potential energy of a particular configuration, $f(x)$ is the probability of x, and λ is a characteristic parameter of each substrate's conformational energy distribution equal to the inverse arithmetic average of all energies sampled for that substrate. The exponential distribution was used because it is functionally equivalent to the Boltzmann distribution sampled by Hyperchem during the Monte Carlo simulation. If d is the characteristic binding distance associated with the potential energy (x) of a sampled configuration, the average characteristic binding distance μ was calculated according to $\mu = \sum(d \times f(x))/\sum f(x)$. Use of the $f(x)$ as a weight factor allowed lower-energy conformations, which are more favorable and thus more probable, to be weighted more heavily in the computation of the average distance than less favorable, higher-energy conformations. In an effort to focus the analyses on the most probable binding configurations, only the distances corresponding to the lowest 5000 potential energy configurations were used in computation of μ.

Results and Discussion

Determination of K_M and k_{CAT}. Measured values of K_M and k_{CAT} are presented in Table 1 alongside corresponding values for the standard error of regression (SE). Items in Table 1 are ranked according to increasing molecular volume, V_M. The significant variation in the measured values, spanning 3 orders of magnitude, prompted us to explore the molecular basis for differences in HRP reactivity.

Relationships between Reaction Rate Constants and Substrate Molecular Structures. Because HRP-mediated reactions involve electron transfer from substrate to heme with simultaneous proton transfer to HIS42, electronic structures of the substrate molecules tend to play a significant role in influencing reactivity.[14,25] We attempted to correlate our measured k_{CAT} values with three electrochemical parameters including E_{HOMO}, E_{LUMO}, and atomic charge on the phenol oxygen. Not surprisingly, no apparent correlation was found between k_{CAT} and either E_{LUMO} or the atomic charge on the phenol oxygen values, as summarized in Table S1 of Supporting Information II. In contrast, a very interesting relationship was found to exist between the k_{CAT} and E_{HOMO}. For 8 of the 15 chemicals studied, k_{CAT} exhibits a strong linear correlation with E_{HOMO} ($R^2 = 0.90$); however, k_{CAT} values for the remaining seven chemicals deviate from the trend with a negative residual such that every value is smaller than anticipated on the basis of the trend (refer to Supporting Information II, Table S1).

Figure 2 depicts substrate degradation rates as a function of E_{HOMO} values. For ease of interpretation, the data have been divided into two groups using different colored symbols: dark squares indicate compounds that fit the trend, and open squares

(39) Wirstam, M.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Am. Chem. Soc.* **1999**, *121*, 10178.
(40) Voet, D.; Voet, J. G. *Biochemistry*, 2nd ed.; John Wiley & Sons Publishing: New York, 1995; p 57.
(41) Kuramochi, H.; Noodleman, L.; Case, D. A. *J. Am. Chem. Soc.* **1997**, *119*, 11442.
(42) Nishihara, T.; Nishikawa, J.; Kanayama, T.; Dakeyama, F.; Saito, K.; Imagawa, M.; Takatori, S.; Kitagawa, Y.; Hori, S.; Utsumi, H. *J. Health Sci.* **2000**, *46*, 282.
(43) *HyperChem Reference Manual*; Hypercube, Inc.: Gainesville, FL, 1996; p 172.

Table 1. Measured ln(k_{CAT}) and K_M Values for 15 Estrogenic Phenols (Columns 1−3) and Simulation-Estimated Binding Distances (Columns 4−7)[a]

	measured parameters			simulation-estimated distances (Å)			
compound	K_M (μM)	ln(k_{CAT}) (s^{-1})	ln(SE k_{CAT}) (s^{-1})	compound I H–HemeO	compound I H–HIS42	compound II H–HemeO	compound II H–HIS42
phenol	614.0	6.71	1.79	5.01	6.61	4.82	7.40
1,4-benzenediol[b,d]	91.8	8.19	−0.04	5.63	7.98	4.52	8.63
1,2,3-benzenetriol	59.1	7.94	0.14	6.13	6.99	5.89	7.50
4-chlorophenol[c]	120.3	7.63	1.49	5.62	6.90	4.29	7.22
4-nitrophenol[b,d]	73.7	6.46	0.14	5.22	7.80	4.51	8.54
4-methoxyphenol[d]	307.9	9.59	0.65	5.27	6.43	5.40	7.35
4-ethylphenol[c]	273.1	9.09	0.14	3.54	6.22	7.12	7.52
4-ethoxyphenol[d]	204.4	9.73	0.92	6.23	5.81	6.16	7.46
2,6-dimethoxyphenol	1188.0	9.51	0.65	5.71	6.80	4.69	7.73
4-*tert*-butylphenol[b,c]	157.6	6.98	−1.66	6.48	8.10	9.62	9.05
4-phenylphenol[b,c]	25.8	7.21	−0.26	4.25	9.22	5.39	10.15
bisphenol A[c]	3.5	9.99	0.30	5.84	5.72	9.07	7.71
4-octylphenol[b,c]	6.0	5.18	1.75	5.01	9.90	8.85	10.78
17β-estradiol[b,c]	12.9	5.22	−1.66	5.94	10.78	11.93	12.40
17α-ethynylestradiol[b,c]	2.6	6.24	1.44	13.74	10.40	9.02	10.61

[a] Distances between the substrate phenolic proton and either the heme-bound oxygen (H−HemeO) or the δN of HIS42's imidazole ring (H−HIS42) are obtained through molecular simulations of enzyme/substrate complexes comprising either compound I (columns 4 and 5) or compound II (columns 6 and 7). [b] Denotes compounds deviating from the E_{HOMO}/ln(k_{CAT}) trend (eq S3 of Supporting Information II). [c] Denotes compounds exhibiting documented estrogenicity.[12,42] [d] Denotes compounds having suspected estrogenicity.[12]

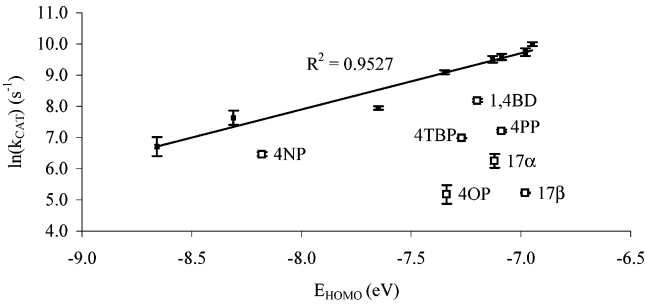

Figure 2. Degradation rate constants as a function of E_{HOMO} for 15 estrogenic phenolic chemicals; error bars represent ±1 SE.

indicate compounds that deviate from it. For ease of reference, data points corresponding to the deviant compounds are also labeled with an abbreviation for the substrate to which they correspond: 4-octylphenol (4OP), 17β-estradiol (17β), 17α-ethynylestradiol (17α), 4-nitrophenol (4NP), 1,4-benzenediol (1,4BD), 4-*tert*-butylphenol (4TBP), and 4-phenylphenol (4PP). At least three interesting observations can be made regarding Figure 2. First, ln(k_{CAT}) generally increases with increasing E_{HOMO}. This is consistent with earlier QSAR studies on HRP-mediated reactions of other phenolic chemicals.[15,22−28] This occurrence is also readily justifiable in that higher E_{HOMO} values are indicative of more favorable electron abstraction and transfer.[23,25] Second, only a portion of the data fits the E_{HOMO} vs k_{CAT} trend. Instead of being randomly scattered about the line, those points that deviate from the trend invariably possess a negative residual. In other words, every compound deviating from the trend exhibits a k_{CAT} value lower than what we would predict by the trend based on its E_{HOMO} value. Third, and more subtle, Figure 2 suggests that molecular size tends to be a factor differentiating between those compounds on and off the line, such that HRP appears to be more capable of mediating degradation of the smaller chemicals.

Quantifying Divergence from the k_{CAT}/E_{HOMO} Trend: Molecular Volume (V_M). For the 15 substrates we measured, the average molecular volume (V_M) for compounds on the line was about 25% less than the average for compounds off the line. This statistically significant ($\alpha = 0.10$) difference in mean molecular volumes leads to the following hypothesis: The potential rate at which a compound may be degraded by HRP is intrinsically determined by its E_{HOMO} value; however, for a compound that is too large or sterically hindered to facilitate optimal binding with the enzyme, the reaction rate will become smaller relative to its maximum potential value.

As a first attempt to evaluate the above hypothesis, we derived a two-parameter correlation equation through linear regression using E_{HOMO} and V_M as independent variables and the reaction rate constant as the dependent variable (eq 1). Correlation

$$\ln k_{CAT} = 19.95 + 1.25 \times E_{HOMO} - 0.006 \times V_M \qquad R^2 = 0.583 \quad (1)$$

between the measured and predicted ln(k_{CAT}) values using eq 1 is depicted graphically in Figure S2 in Supporting Information III. In comparison to Figure 2, addition of V_M as a second predictor unites the compounds on and off the trend into a single group reasonably well fit by eq 1. Still, the somewhat low correlation coefficient suggests that V_M, a measure of a substrate's bulkiness, is insufficient to explain the divergence from the k_{CAT}/E_{HOMO} trend by itself. We thus believe that substrate V_M, albeit indirectly related to enzyme/substrate binding interactions, does not capture completely the way in which enzyme/substrate binding facilitates optimal electron transfer.

Quantifying Divergence from the k_{CAT}/E_{HOMO} Trend: Enzyme−Substrate Binding Conformation. As described earlier, previous studies on HRP/substrate bindings indicate that the magnitude of two binding distances may play a critical role in modulating the electron transfers, i.e., the distances from the substrate's phenolic proton to either the heme-bound oxygen[25,44] or to the δN of HIS42's imidazole ring.[32−34] Using computational simulations described in the Experimental Section, we achieved an energy-weighted average value for each of the two binding distances described above in both compounds I and II, which are listed in columns 4−7 of Table 1. Substrates in Table 1 are ranked by increasing molecular volume, and the superscript b denotes the seven compounds that deviate from

(44) Sakurada, J.; Takahashi, S.; Hosoya, T. *J. Biol. Chem.* **1986**, *261*, 0657.

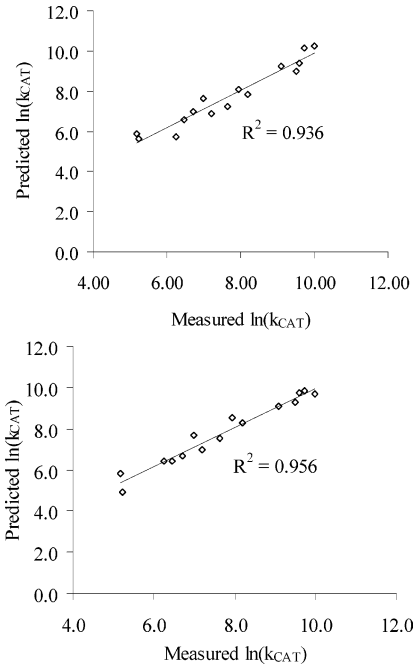

Figure 3. Correlations between measured and predicted values of ln(k_{CAT}) using eq 2 for compound I (top) and eq 3 for compound II (bottom).

the E_{HOMO}/k_{CAT} trend. Values of trend residuals for each substrate are outlined in Table S1 of Supporting Information II.

It became immediately obvious that there is no significant correlation between the residual from the E_{HOMO}/k_{CAT} trend and the H−HemeO distance for either compound I ($R^2 = 0.12$) or compound II ($R^2 = 0.42$). In contrast, the deviation from the E_{HOMO}/k_{CAT} trend is strongly correlated to the H−HIS42 distance for both compound I ($R^2 = 0.92$) and compound II ($R^2 = 0.96$). Such significant correlation, in conjunction with the poor H−HemeO distance correlation, substantiates that HIS42 plays an important role in the electron-transfer process. HIS42 may serve as the point of contact for electron transfer or may facilitate the electron transfer between the substrate and the heme through deprotonating the substrate. Further, both correlations are negative, indicating that increased binding distance is associated with increased rate retardation relative to the potential rate determined by a substrate's intrinsic E_{HOMO} value. By using the H−HIS42 distances in compound I (μ^{I}_{H-HIS}) or compound II complexes (μ^{II}_{H-HIS}) and E_{HOMO} as independent variables, we developed two-parameter linear regression equations for ln(k_{CAT}), eqs 2 and 3. Figure 3 illustrates correlations between the measured and predicted ln(k_{CAT}) values using eqs 2 and 3.

$$\ln k_{CAT} = 25.39 + 1.43 \times E_{HOMO} - 0.91 \times \mu^{I}_{H-HIS} \qquad R^2 = 0.936 \quad (2)$$

$$\ln k_{CAT} = 30.86 + 1.94 \times E_{HOMO} - 1.00 \times \mu^{II}_{H-HIS} \qquad R^2 = 0.956 \quad (3)$$

As mentioned earlier, the reaction mediated by compound II is much slower than that mediated by compound I and thus determines the overall rates of the HRP catalytic cycle (refer to Figure 1). As such, it follows that the overall reaction rates measured in this study were supposed to be predominately influenced by the substrate binding to compound II. That the apparent rates were also well correlated to substrate binding to compound I may suggest that enzyme/substrate bindings with compounds I and II are correlated with each other, at least with respect to characteristic H−HIS42 distance. This is reasonable insofar as compounds I and II share similar amino acid conformations in the immediate vicinity of the active site and differ only slightly in charge distribution on the heme iron and iron-bound oxygen atoms themselves. An interesting observation that can be made upon consideration of Table 1 is related to binding distances for substrates that fit the E_{HOMO}/k_{CAT} trend: in compounds I and II, all substrates that fit the trend share similar H−HIS42 binding distances. For compound I, the average H−HIS distance (μ^{I}_{H-HIS}) is 6.44 ± 0.17 Å (SE). In contrast, the average H−HIS distance for compound II (μ^{II}_{H-HIS}) is slightly larger at 7.49 ± 0.06 Å (SE). For those chemicals that deviate from the trend, H−HIS42 distances are also generally greater for compound II complexes when compared to those for compound I complexes. A paired t-test was performed using binding distances μ^{I}_{H-HIS} and μ^{II}_{H-HIS} for a given substrate. The result indicates that the average difference between binding distances ($\mu^{II}_{H-HIS} - \mu^{I}_{H-HIS}$) is statistically significant (p value < 0.001) with a value of 0.96 ± 0.13 Å (SE). This increase in binding distance from compound I to compound II may contribute to the generally slower reaction rate mediated by compound II relative to compound I.

Comparison of eqs 1, 2, and 3 indicates that the H−HIS42 distance has much greater predictive power relative to V_M in quantifying the impact of enzyme/substrate conformation on apparent reaction rate. The subtle differentiation between molecular volume and binding distance is exemplified by consideration of bisphenol A binding. Although this compound has the fourth largest molecular volume among the 15 chemicals studied (Supporting Information II, Table S1), its reaction rate constant fits the E_{HOMO}/k_{CAT} trend quite well (residual = 0.20). In contrast, 4-phenylphenol, a chemical of similarly large molecular volume, has a reaction rate constant much smaller than that predicted by the trend (residual = −2.33). Our simulations show that the H−HIS42 distances for bisphenol A in compound I and II complexes are 5.72 and 7.71 Å, respectively. These distances are indeed much smaller than those for 4-phenylphenol (9.22 and 10.15 Å in compound I and II complexes, respectively). The bisphenol A characteristic binding distances are also similar to those of the other chemicals that fit the E_{HOMO}/k_{CAT} trend. The ability of bisphenol A to make closer contact with HIS42 might arise from such factors as its bilateral symmetry and linearity in shape. Because these molecular characteristics are not accurately encapsulated in the measurement of molecular volume, V_M is a less powerful predictor of degradation rate compared to binding conformation.

Conclusion

This work demonstrates that the rates of HRP-mediated reactions of phenols are well correlated to two parameters: one describing the intrinsic redox reactivity of the substrates (E_{HOMO}) and the other describing the enzyme/substrate binding conformation (as quantified using the H−HIS42 distance). This validates the importance of HIS42 in electron transfer mediated by HRP. The two-parameter QSAR equations comprising E_{HOMO} and H−HIS42 distances in compound I or compound II

complexes as independent variables yield high correlation coefficients ($R^2 = 0.94$ and 0.95, respectively). This implies that the enzyme/substrate bindings with compounds I and II are correlated with each other. Further, this QSAR study indicates that HRP reactivity is significantly impacted by the enzyme/substrate binding conformation with respect to optimal H−HIS42 distance. This suggests that proteomic modifications of HRP, strategically designed to decrease the H−HIS42 distance for selected substrates of concern, may be useful in enhancing HRP reactivity toward recalcitrant phenolic contaminants and improving degradation efficiencies in engineering applications. Such proteomic modifications may be designed and screened using the QSAR equations developed in this study.

Acknowledgment. This research was financed by Research Grant P42ES04911-14 from the National Institutes for Environmental and Health Sciences. L.M.C. is thankful for support of a National Science Foundation Graduate Research Fellowship.

Supporting Information Available: Charge distribution for selected subsets of enzyme intermediates (I); molecular descriptors of substrates (II); correlation between measured and predicted degradation rates using eq 1 (III); expanded reference (IV). This material is available free of charge via the Internet at http://pubs.acs.org.

JA057430F