ACS Catalysis

Research Article

pubs.acs.org/acscatalysis

Relationship between Enzyme/Substrate Properties and Enzyme Efficiency in Hydrolases

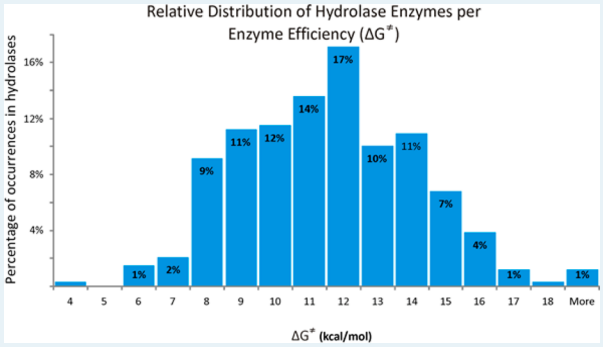
Sérgio F. Sousa,[†] Maria J. Ramos,[†] Carmay Lim,[*,‡,§] and Pedro A. Fernandes[*,†]

[†]UCIBIO, REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal

[‡]Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan

[§]Department of Chemistry, National Tsing Hua University, Hsinchu 300, Taiwan

ABSTRACT: Hydrolase enzymes are involved in breaking different chemical bonds in diverse substrates of different sizes and complexity such as proteins, carbohydrates, lipids, and nucleic acids. This work presents a systematic analysis of the kinetic, structural, and biological information on hydrolases, taking into account the presence of different cofactors, the number of chains (i.e., oligomerization state), the number of amino acid residues, and the physicochemical properties of the substrates. Specific trends related to the catalytic activity of this large and rather diverse class of enzymes, including activation free energies, binding free energies, and enzyme efficiencies are revealed and rationalized. The results show that despite the diversity of hydrolases, their substrates, and reactions, hydrolases have quite characteristic activation free energies, substrate binding free energies, and enzyme efficiencies. Different subclasses of hydrolases employ different strategies to achieve these values, including the use of cofactors, different numbers of chains, and numbers of amino acid residues. The large structural and physicochemical diversity of the substrates acted by hydrolases is heavily neutralized by the hydrolases, resulting in activation free energies, binding free energies and enzyme efficiencies that are, in general, nearly independent of the diversity of the substrates.



KEYWORDS: *activation free energy, catalysis, catalytic power, rate enhancement, enzyme evolution*

INTRODUCTION

Enzymes play an essential biological role in performing and controlling an important share of the chemical processes occurring in life. Understanding how enzymes can achieve their tasks is a problem of widespread interest and importance, both from a fundamental and practical point of view.[1] Over the past few decades, several hypotheses have been put forth to account for the observed enzyme rate enhancements,[1b,2] which range from 10^8 to 10^{27} M^{-1}.[1a] These include (i) reduction of orientational entropy,[3] (ii) orbital steering,[4] (iii) steric tension[5] or desolvation[6] of the reactants, (iv) correlated structural fluctuations,[2f] (v) electrostatic preorganization of the active site,[2b] (vi) tuning of the acid/base pK_a, (vii) low barrier hydrogen bonds,[7] (viii) change in the reaction mechanism through participation of aa residues or cofactors,[1a] (ix) nonequilibrium specific vibrations,[8] and (x) tunneling.[2e]

Today, almost 70 years after the proposal by Linus Pauling that an enzyme can lower the activation energy because it has higher affinity for the transition state than the substrate,[9] there is still no full consensus regarding how all the different enzymes accomplish this differential binding to achieve their catalytic prowess, although electrostatic preorganization seem to be dominant in most enzymes studied to date.[2h,10] Elucidating the relative contributions of the various features responsible for the enormous catalytic power of enzymes remains a key challenge.[10a,11] This challenge stems from the vast diversity of enzyme structures, functions, mechanisms, as well as their size, molecular complexity, unusual chemistry, and wide time range of molecular catalytic events. Hydrolases, which catalyze the hydrolysis of chemical bonds, illustrates such diversity with more than 1200 different hydrolases of various structures catalyzing the hydrolysis of many different substrates.[12] Their structures comprise 38% of all the enzyme structures in the January 2015 release of the Protein Data Bank.[13]

Hydrolases are characterized by enzyme commission (EC) number EC3 and are classified into 13 subclasses according to the specific bonds cleaved. These 13 subclasses comprise esterases (EC3.1), including nucleases, phosphodiestereases, lipases, and phosphatases, which catalyze hydrolysis of ester bonds; glycosylases (EC3.2) which hydrolyze sugars;[14] ether, thioether ad trialkylsulfonium hydrolases (EC3.3) that break up ether bonds; proteases/peptidases (EC3.4) such as amino-peptidases, carboxypeptidases, and endopeptidases, which hydrolyze peptidic C−N bonds; hydrolases that hydrolyze

Received: May 5, 2015
Revised: July 13, 2015
Published: August 24, 2015

© 2015 American Chemical Society

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887

nonpeptidic C−N bonds in amides, amidines, and nitriles (EC3.5); hydrolases acting on acid anhydrides including nucleoside di and triphosphates and sulfonyl-containing anhydrides (EC3.6); hydrolases acting on C−C bonds (EC3.7),[15] C−halide bonds (EC3.8), P−N bonds (EC3.9), S−N bonds (EC3.10), C−P bonds (EC3.11), S−S bonds (EC3.12), and C−S bonds (EC3.13). Clearly, hydrolases act on very different bonds with markedly different bond strengths and stabilities.[16] In fact, the rates of uncatalyzed hydrolysis reactions at 25 °C span ∼17 orders of magnitude with solution activation free energies ranging from 21 to 44 kcal/mol.[16a] Hence, the challenge posed by Nature to the diverse hydrolases can be quite different. This feature and the fact that all hydrolases use a water molecule to perform their function makes this class of enzymes an appealing test case for statistical analyses to elucidate the astonishing power of enzymes.

In this work, we present a detailed analysis of the features behind the catalytic activity of this vast and chemically diverse class of enzymes. We dissect the molecular basis underlying the activation free energies, binding free energies, and enzyme efficiencies of a curated data set of 339 reactions catalyzed by hydrolases. We then perform statistical tests to identify the statistical significant similarities/differences in the catalytic activity of this large and rather diverse class of enzymes.

METHODOLOGY

CREATING THE HYDROLASE DATABASE

Wild-type hydrolases with well-defined four EC numbers and recorded temperature and pH were extracted from the BRENDA enzyme database.[17] Enzymes lacking turnover and K_M values and/or with unspecified experimental conditions (temperature and pH) as well as mutant species were excluded. This resulted in a total of 887 hydrolases with consistent turnover and K_M entries corresponding to the same experimental conditions.

Because an enzyme could react with non-native substrates in artificial conditions typically with lower reaction rates, it is critical to verify that the enzyme parameters from BRENDA correspond to *native* substrates. Hence, each substrate in the initial data set was verified according to known enzyme−substrate specificity determinants and current knowledge using the enzyme annotations in the ExPASy resource portal.[18] Whenever necessary, substrate specificity information was further checked or complemented with information available from the most recent literature for each enzyme. The cleaned-up database contains 339 hydrolases with confirmed physiological substrates and associated enzyme parameters.

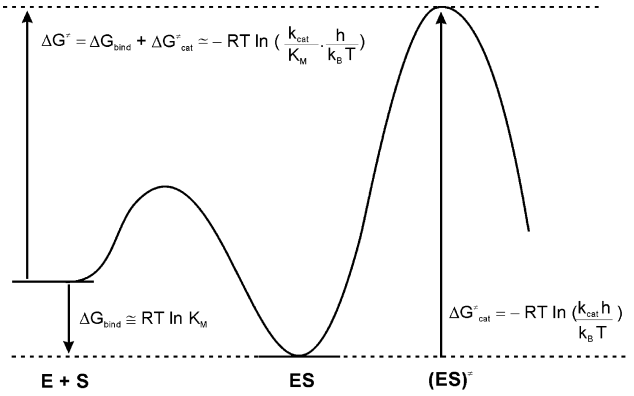
Enzyme Properties Analyzed. Because the turnover number is equal to the first-order rate constant k_{cat} when the enzyme is saturated with substrate, regardless of the rate-limiting step, it was used to approximate the k_{cat}. The enzyme's catalytic efficiency was estimated by k_{cat}/K_M. The K_M, k_{cat}, and k_{cat}/K_M, values and the corresponding temperatures were used to determine ΔG_{bind}, $\Delta G^{\neq}_{cat}$, and $\Delta G^{\neq}$ (see Scheme 1):[19]

Note that for hydrolases that perform reversible reactions, the associated kinetics would be best described by Haldane kinetics (quasi steady-state approximation). For these enzymes, the binding free energies from the K_M values should be taken as an approximation to the actual binding free energies.

In addition to the kinetic parameters, each enzyme entry was further manually complemented with structural and biological information from BRENDA[17] and UniProtKB/Swiss-Prot databases.[20] Features added included the cellular location of the specific isoform, the number of chains (i.e., oligomerization state), the number of aa residues, and the type of cofactor(s). Each enzyme entry was further classified as acting on a single specific substrate or on more than one substrate or class of substrates.

In addition to the information present in BRENDA and the UniProtKB/Swiss-Prot databases, the consistency of the information for each entry and species such as substrate preference, number of amino acid residues, number of chains in the native active enzyme, native metal cofactor was confirmed from (i) the corresponding references in the literature, (ii) available structures in the Protein Data Bank and their references, and (iii) the amino acid sequences of the enzyme in the UniProtKB database. Consistency of the data reported was also manually checked taking into consideration the similarity of the species; that is, a hydrolase enzyme from a particular mammal is not expected to have drastically different metal dependence or number of amino acid residues from the same enzyme in other related mammals. These procedures were performed to validate the data, correct inconsistencies, and minimize possible errors in the data set.

Scheme 1. Schematic Description of the Energetics of Enzymatic Reactions, Illustrating the Relationship between K_M, k_{cat}, and k_{cat}/K_M Values and ΔG_{bind}, $\Delta G^{\neq}_{cat}$, and $\Delta G^{\neq}$[a]



[a] $\Delta G^{\neq}$ gives a measure of the catalytic efficiency (it is related to the second-order rate constant, relevant for enzymes and substrates at physiological concentration), ΔG_{bind} is the binding free energy, and $\Delta G^{\neq}_{cat}$ is the apparent activation free energy for the first-order reaction when the substrate is present at saturating concentrations.

SUBSTRATE PROPERTIES ANALYZED

For each substrate, Open Babel[21] was used to generate structures at the corresponding pH. The protonation state of ionizable residues in each structure was visually inspected and corrected as required. To generate physical and structural properties for each substrate, the Molecular Operating Environment (MOE) was used to compute the following properties for each substrate; viz., molecular weight, volume, number of atoms, diameter (defined by the largest value in the distance matrix involving all the atoms in the molecule), number of rotatable bonds, number of hydrogen bond donors and acceptors, log P (where P is the partition coefficient between octanol and water), and log S (where S is the solubility), as well as accessible surface area (ASA) considering a probe radius of 1.4 Å. The Molecular Operating Environment

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887

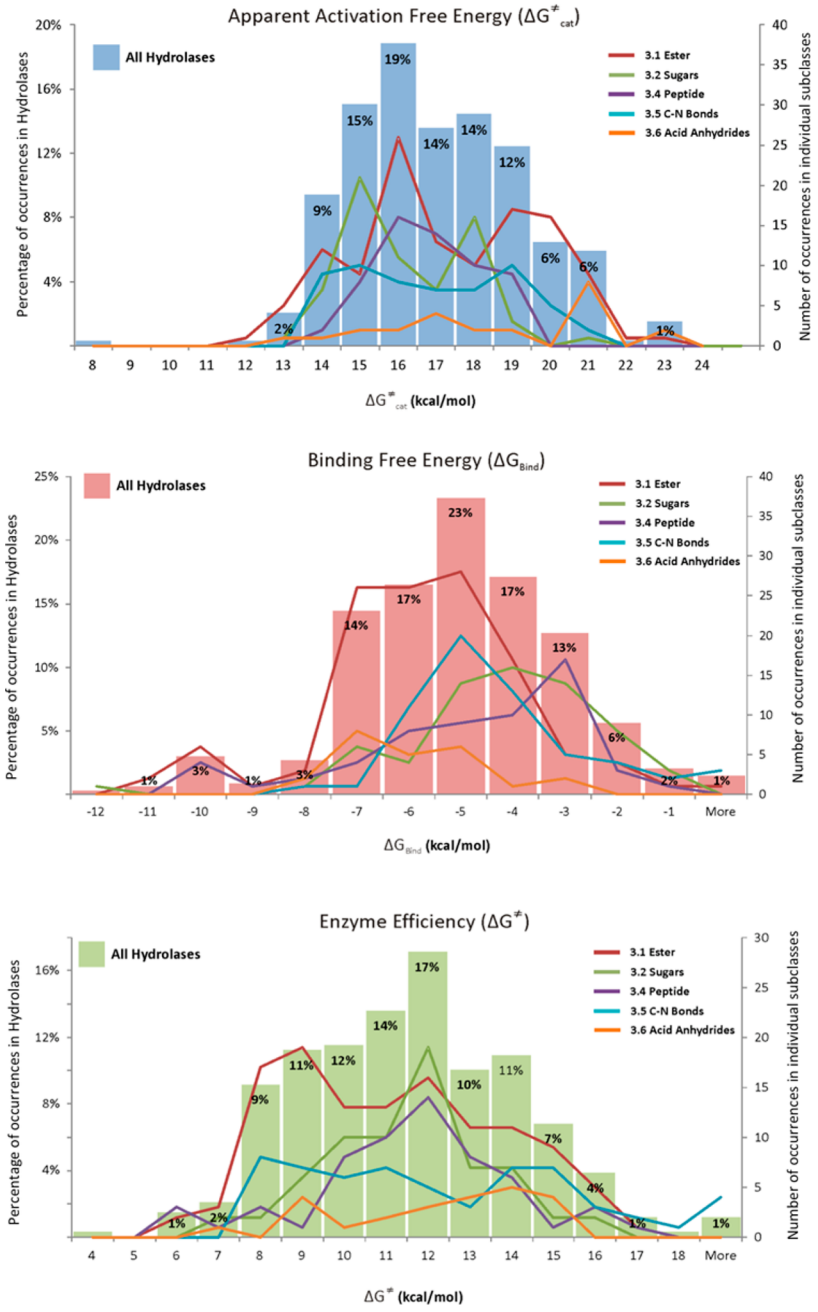


Figure 1. Distribution of (a) ΔG_{bind}, (b) $\Delta G^{\neq}_{cat}$, and (c) $\Delta G^{\neq}$ among hydrolases. The % occurrence frequencies for all hydrolases are shown as histograms, whereas those for subclasses with >20 entries are shown as curves in different colors.

used the partial charges on atom i to decompose the ASA into positive ($q_i > 0$), negative ($q_i < 0$), polar ($|q_i| > 0.2$) and hydrophobic ($|q_i| < 0.2$) fractions.

Statistical Analysis. For each enzyme subclass with more than 20 entries, the average value of each enzyme parameter as a function of a given enzyme or substrate property was computed. To determine if this is statistically different from the respective average value of each of the other subclasses, a two-tailed t-test was performed. The two-tailed t-test gives the probability p that two data sets are not statistically different, taking into account the size of each data set, the distribution of the corresponding values, and its average value. The confidence level that the mean values derived from two data sets are statistically different is given by $(1 - p) \times 100$. When the probability that two data sets are not statistically different was computed to be less than 5%, the average values derived from the two data sets were deemed to be statistically different.

RESULTS

We first report the variations in the substrate-binding free energies (ΔG_{bind}), apparent activation free energy ($\Delta G^{\neq}_{cat}$), and the catalytic efficiency ($\Delta G^{\neq}$) for those hydrolases with >20 entries in each subclass and their correlations. Consistent ΔG_{bind}, $\Delta G^{\neq}_{cat}$, and $\Delta G^{\neq}$ values were unavailable for hydrolases acting on C−C bonds (EC3.7), P−N bonds (EC3.9), S−N bonds (EC3.10), S−S bonds (EC3.12), and C−S bonds (EC3.13), whereas only three sets of free energies were found for each subclass of hydrolases acting on ethers (EC3.3), halide bonds (EC3.8), and C−P bonds (EC3.11), and thus, analyses were not performed for these hydrolase subclasses. We reveal

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887

how these enzyme parameters depend on properties of the enzyme and substrate. We then discuss those mean values derived from two data sets that are found to be statistically different according to two-tailed *t*-tests (see Methods), and we list the corresponding confidence level that the two average values are different. When the confidence level is ≤35%, the two averages are assumed to be similar.

Variations in the Enzyme Parameters. ΔG_{bind}. Hydrolases bind an incredibly diverse set of substrates, which can be as small as diphosphate and acetamide or as large as starch, angiotensin I, neurotensin, polynucleotides, and polysaccharides. Yet, the substrate-binding free energies (ΔG_{bind}) fall within a relatively limited range (−3 to −7 kcal/mol) for 84% of the hydrolases (Figure 1a) with an average ΔG_{bind} of −5.5 ± 2.0 kcal/mol for all hydrolases (Table 1). Zhang and Houk[1a] found binding free energies for diverse types of host−guest complexes spanning from +3 to −15 kcal/mol. Likewise, Smith and co-workers[22] observed a similar binding free energy variation for a collection of protein−ligand complexes from the Binding MOAD database[23] with most values ranging from −3 kcal/mol to −15 kcal/mol and a few as large as −20 kcal/mol. So the ΔG_{bind} variation for hydrolases can be regarded as relatively narrow, despite reflecting a ∼1000 times difference in binding affinity.

Table 1. Average Values of ΔG_{bind}, $\Delta G^{\neq}_{cat}$, and $\Delta G^{\neq}$ for Hydrolases with >20 Entries in Each Subclass

subclass/ substrate	no. of enzymes	$\langle\Delta G_{bind}\rangle$ (kcal/mol)	$\langle\Delta G^{\neq}_{cat}\rangle$ (kcal/mol)	$\langle\Delta G^{\neq}\rangle$ (kcal/mol)
3.1 esters	120	−6.2 ± 2.0	16.8 ± 2.4	10.6 ± 2.5
3.2 sugars	67	−4.7 ± 1.8	15.7 ± 1.6	11.0 ± 1.9
3.4 peptides	59	−5.3 ± 2.1	16.4 ± 1.4	11.1 ± 2.4
3.5 C−N bonds	60	−4.7 ± 1.9	16.6 ± 2.3	11.9 ± 3.3
3.6 acid anhydrides	24	−6.3 ± 1.3	18.0 ± 2.3	11.6 ± 2.4
all hydrolases	**339**	**−5.5 ± 2.0**	**16.6 ± 2.2**	**11.1 ± 2.6**

$\Delta G^{\neq}_{cat}$. The apparent activation free energy $\Delta G^{\neq}_{cat}$ for the first-order reaction (i.e., when the substrate is present at saturating concentrations) catalyzed by 95% of the hydrolases ranges from 14 to 21 kcal/mol (Figure 1b). The average $\Delta G^{\neq}_{cat}$ for all 339 hydrolases is 16.6 ± 2.2 kcal/mol (Table 1). The standard deviation of 2.2 kcal/mol is remarkably small considering the large structural diversity of the corresponding enzymes, substrates, and the wide range of bonds being broken during catalysis. Only EC3.2 glycosylases have a significantly smaller mean $\Delta G^{\neq}_{cat}$ (15.7 ± 1.6) than hydrolases in other subclasses within a 95% confidence level.

$\Delta G^{\neq}$. The catalytic efficiency, as measured by the difference in free energy between the rate-limiting TS and the separated reactants, ranges from 8 to 14 kcal/mol for 83% of the hydrolases (Figure 1c) with an average $\Delta G^{\neq}$ of 11.1 ± 2.6 kcal/mol for all hydrolases (Table 1). The mean $\Delta G^{\neq}$ values for the subclasses in Table 1 appear more similar than the activation or binding free energies, as they were not different within a confidence level of 95% except for the comparison between the mean $\Delta G^{\neq}$ values of EC3.1 (10.6 kcal/mol) and EC3.5 (11.9 kcal/mol) hydrolases (99.2% confidence level). Compared to the EC3.5 hydrolases that hydrolyze nonpeptidic C−N bonds, EC3.1 esterases appear to be more catalytic efficient because they tend to bind substrates better (−6.2 vs −4.7 kcal/mol, 99.99% confidence level) but exhibit similar mean $\Delta G^{\neq}_{cat}$ (16.8 vs 16.6 kcal/mol).

Correlation between $\Delta G^{\neq}$ and $\Delta G^{\neq}_{cat}$ or ΔG_{bind}. Because $\Delta G^{\neq}$ was estimated as a sum of $\Delta G^{\neq}_{cat}$ and ΔG_{bind} (see Scheme 1), $\Delta G^{\neq}_{cat}$ and ΔG_{bind} can be expected to correlate with $\Delta G^{\neq}$, but which of these two quantities has more influence on the catalytic efficiency of hydrolases belonging to different subclasses? To address this, we correlated $\Delta G^{\neq}$ with $\Delta G^{\neq}_{cat}$ or ΔG_{bind} for the different subclasses of hydrolases and computed the corresponding slope and r^2 values (Table 2). For hydrolases acting on esters (EC3.1), nonpeptidic C−N bonds (EC3.5), and acid anhydrides (EC3.6), the catalytic efficiency appears more dependent on the activation free energy (Figure 2a) than on the binding free energy: $\Delta G^{\neq}$ varies more with $\Delta G^{\neq}_{cat}$ (slopes of 0.64, 0.58, and 1.04) than with ΔG_{bind} (slopes 0.35, 0.41, and −0.04). On the other hand, for hydrolases acting on sugars (EC3.2) and peptides (EC3.4), the catalytic efficiency seems more sensitive to the substrate binding free energy (Figure 2b) than to the activation free energy: the $\Delta G^{\neq}$ variation correlates better with ΔG_{bind} (slopes 0.62 and 0.72) than with $\Delta G^{\neq}_{cat}$ (slopes 0.37 and 0.27). Differences in the substrate sizes may explain why the catalytic efficiency of EC3.4 hydrolases depends more on peptidase binding, whereas that of EC3.5 hydrolases depends more on their activation free energies, even though both types of enzymes hydrolyze C−N bonds: The average molecular weight of EC3.4 hydrolase substrates (550 g/mol) is roughly twice that of EC3.5 hydrolase substrates (242 g/mol). Interestingly, $\Delta G^{\neq}_{cat}$ do not seem to correlate with ΔG_{bind} for hydrolases acting on C−N bonds ($r^2 \sim 0$) and only weakly for the other hydrolase subclasses ($0.1 < r^2 < 0.3$).

Table 2. Correlation between $\Delta G^{\neq}$ and $\Delta G^{\neq}_{cat}$ or ΔG_{bind}

		$\Delta G^{\neq}_{cat}$ vs $\Delta G^{\neq}$		ΔG_{bind} vs $\Delta G^{\neq}$		ΔG_{bind} vs $\Delta G^{\neq}_{cat}$	
subclass/substrate	no. of enzymes	slope	r^2	slope	r^2	slope	r^2
3.1 esters	120	0.64	0.47	0.35	0.21	−0.27	0.11
3.2 sugars	67	0.37	0.20	0.62	0.40	−0.47	0.17
3.4 peptides	59	0.27	0.22	0.72	0.65	−0.20	0.02
3.5 C−N bonds	60	0.58	0.71	0.41	0.55	0.21	0.07
3.6 acid anhydrides	24	1.04	0.78	−0.04	0.01	−0.25	0.28

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887

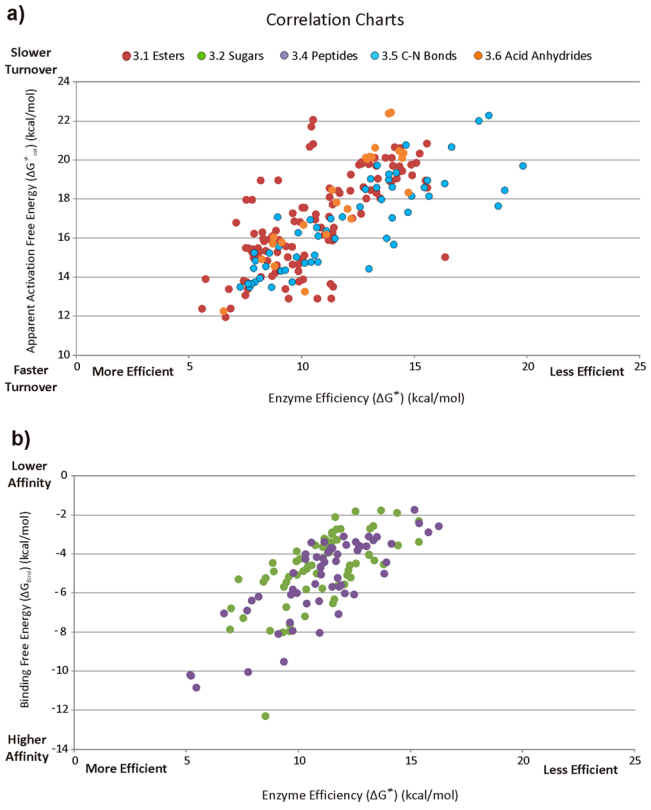


Figure 2. Correlation between $\Delta G^{\neq}$ and (a) $\Delta G^{\neq}_{cat}$ or (b) ΔG_{bind} for the different hydrolase subclasses that exhibit Pearson correlation coefficients >0.5.

Dependence of Enzyme Parameters on Cofactors. Cofactors play important roles in enzymatic catalysis and in the structural stabilization of proteins;[24] hence, they can be expected to increase the catalytic efficiency. Among the 339 hydrolases analyzed herein, roughly half (175) employ metal or organic cofactors, whereas the remaining enzymes do not seem to employ any cofactor, albeit information on such native cofactors could be missing. Hydrolases that employ cofactors exhibit the same range of $\Delta G^{\neq}_{cat}$, ΔG_{bind}, and $\Delta G^{\neq}$ values as those that do not use cofactors (Figure 3a), but the latter have on average higher activation free energies (16.9 vs 16.3 kcal/mol, 98.8% confidence level), less favorable binding free energies (−5.2 vs −5.8 kcal/mol, 99.8% confidence level), and are thus less catalytic proficient (11.7 vs 10.5 kcal/mol, 99.999% confidence level). To verify that cofactors generally enhance the catalytic efficiency of these enzymes, we computed the difference between the percentage frequency of a given free energy for hydrolases without cofactors and that for hydrolases that employ cofactors. The results in Figure 3b show that hydrolases that employ cofactors are prevalent among the hydrolases that bind substrates relatively tightly (ΔG_{bind} between −6 and −12 kcal/mol) with activation free energies $\Delta G^{\neq}_{cat}$ < 18 kcal/mol and catalytic efficiencies $\Delta G^{\neq}$ < 12 kcal/mol, whereas those that do not use cofactors are prevalent among the hydrolases that bind substrates weakly (ΔG_{bind} between −1 and −6 kcal/mol) with higher $\Delta G^{\neq}_{cat}$ (18−21 kcal/mol) and $\Delta G\neq$ (12−18 kcal/mol).

Dependence of Enzyme Parameters on the Cofactor Type. Hydrolases employ various metal cofactors such as Zn^{2+}, Mg^{2+}, Ca^{2+}, and Mn^{2+}, which are embedded in mononuclear or binuclear binding sites.[25] What is the advantage (if any) of a binuclear compared to a mononuclear metal site in enzyme catalysis? Only Zn^{2+}, Mg^{2+}, and Ca^{2+} hydrolases have sufficient (>20) consistent enzyme parameters. These were divided into two groups according to whether their metal-binding sites are mononuclear or binuclear; in each group with >20 entries, the mean ΔG_{bind}, $\Delta G^{\neq}_{cat}$, and $\Delta G^{\neq}$ were computed (Table 3). Only Zn^{2+} hydrolases have >20 mononuclear and binuclear Zn-sites, whereas Mg^{2+} hydrolases have only four confirmed binuclear sites, whereas all Ca^{2+} hydrolases with consistent enzyme

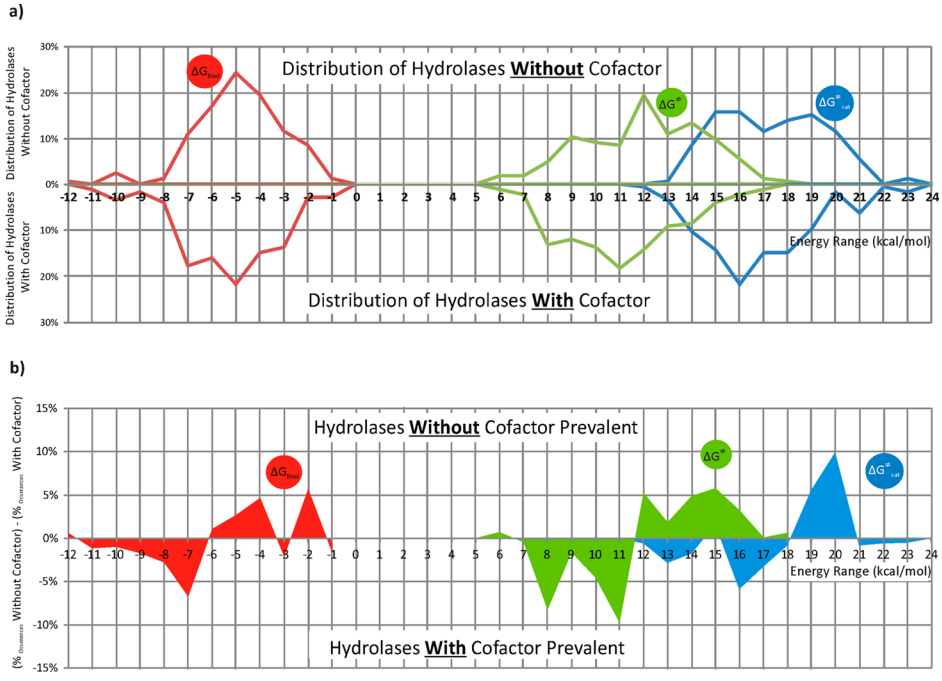


Figure 3. (a) Percentage occurrence frequencies of $\Delta G^{\neq}_{cat}$, ΔG_{bind}, and $\Delta G^{\neq}$ of hydrolases with and without cofactors. (b) Percentage occurrence frequency of a given $\Delta G^{\neq}_{cat}$, ΔG_{bind}, or $\Delta G^{\neq}$ of hydrolases that do not employ cofactors minus that of hydrolases that do employ cofactors.

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887

Table 3. Average $\Delta G^{\neq}_{cat}$, ΔG_{bind}, and $\Delta G^{\neq}$ for Hydrolases Containing Zn^{2+}, Mg^{2+}, and Ca^{2+}

cofactor	no. of enzymes	average $\Delta G^{\neq}$ (kcal/mol)	average $\Delta G^{\neq}_{cat}$ (kcal/mol)	average ΔG_{bind} (kcal/mol)
Zn^{2+} mononuclear	26	11.1 ± 2.3	16.9 ± 2.1	−5.8 ± 1.7
Zn^{2+} binuclear	43	10.1 ± 2.1	15.6 ± 1.3	−5.5 ± 1.6
Mg^{2+} mononuclear	22	10.1 ± 2.1	17.1 ± 1.8	−7.0 ± 1.0
Ca^{2+} mononuclear	29	10.2 ± 2.3	16.1 ± 1.4	−5.9 ± 2.3

parameters contain mononuclear sites. The presence of a second Zn^{2+} seems to enhance catalytic efficiency: hydrolases with binuclear Zn-sites exhibit lower mean $\Delta G^{\neq}$ than those with mononuclear Zn-sites (10.1 vs 11.1 kcal/mol, 96.0% confidence level). The second Zn^{2+} affects the activation barrier (lowering the mean $\Delta G^{\neq}_{cat}$ for mononuclear sites from 16.9 to 15.6 kcal/mol, 99.2% confidence level) rather than improving substrate binding.

Which metal cofactor is more effective as a coenzyme and does it enhance catalytic efficiency through $\Delta G^{\neq}_{cat}$ or ΔG_{bind}? To address this question, we compared the enzyme parameters of mononuclear Zn^{2+}, Mg^{2+}, and Ca^{2+} hydrolases, for which there are >20 entries. For hydrolases with mononuclear sites, Mg^{2+} and Ca^{2+} appear to be more effective than Zn^{2+} as a coenzyme: the mean $\Delta G^{\neq}$ for mononuclear Mg^{2+} or Ca^{2+} hydrolases is smaller than that the corresponding Zn^{2+} enzymes (10 vs 11 kcal/mol, 91.0% confidence level). However, Mg^{2+} enhances substrate binding, whereas Ca^{2+} apparently reduces the activation barrier in improving catalytic efficiency over Zn^{2+}: Compared to the mononuclear Zn^{2+} enzymes, mononuclear Mg^{2+} hydrolases bind substrates more favorably (−7.0 vs −5.8 kcal/mol, 99.5% confidence level), whereas mononuclear Ca^{2+} hydrolases exhibit a lower mean $\Delta G^{\neq}_{cat}$ (16.1 vs 16.9 kcal/mol, 86% confidence level). Interestingly, Mg^{2+} and Ca^{2+} hydrolases achieve their catalytic efficiency (10.1 vs 10.2 kcal/mol) in different ways: Compared to the Ca^{2+} enzymes, Mg^{2+} hydrolases bind substrates more favorably (−7.0 vs −5.9 kcal/mol, 96.9% confidence level), but have larger mean $\Delta G^{\neq}_{cat}$ (17.1 vs 16.1 kcal/mol, 96.5% confidence level).

Dependence of Enzyme Parameters on the Oligomeric State. Hydrolases can function as monomers, dimers, trimers, tetramers, or hexamers. Does the oligomeric state help to bind substrate and/or lower the activation barrier? To address this question, we computed the mean ΔG_{bind}, $\Delta G^{\neq}_{cat}$, and $\Delta G^{\neq}$ and corresponding standard deviations of monomers, dimers, or tetramers (Table 4), but not of trimers, pentamers, and hexamers for which there were too few consistent enzyme parameters available. Oligomerization does not seem to help substrate binding or barrier reduction. Compared to monomers, dimers and tetramers seem to be less efficient with a higher mean $\Delta G^{\neq}$ (~11.5 vs 10.4 kcal/mol, 99.97% confidence level) and higher mean $\Delta G^{\neq}_{cat}$ (~17 vs 16 kcal/mol, 99.9% confidence level).

Table 4. Mean $\Delta G^{\neq}_{cat}$, ΔG_{bind}, and $\Delta G^{\neq}$ as a Function of the Hydrolase Oligomeric State

oligomeric state	no. of enzymes	average $\Delta G^{\neq}$ (kcal/mol)	average $\Delta G^{\neq}_{cat}$ (kcal/mol)	average ΔG_{bind} (kcal/mol)
monomer	90	10.4 ± 2.1	16.1 ± 1.9	−5.7 ± 2.2
dimer	76	11.4 ± 3.3	16.6 ± 2.4	−5.2 ± 2.5
tetramer	94	11.6 ± 2.4	17.3 ± 2.1	−5.7 ± 1.6
homo-oligomer	136	11.2 ± 2.6	16.8 ± 2.3	−5.6 ± 1.6
hetero-oligomer	28	11.4 ± 2.3	17.8 ± 2.7	−6.4 ± 2.7

To determine if homo-oligomeric enzymes are more competent in substrate binding or in lowering the activation barrier than hetero-oligomeric ones, we compared the mean ΔG_{bind}, $\Delta G^{\neq}_{cat}$, and $\Delta G^{\neq}$ for 136 homo-oligomeric hydrolases with 28 hetero-oligomeric ones. The confidence level that the mean $\Delta G^{\neq}$ values for homo- and hetero-oligomeric hydrolases are different (11.2 vs 11.4 kcal/mol) is only 35%, indicating that homo-oligomeric hydrolases may not be more catalytic proficient than hetero-oligomeric hydrolases. Although homo-oligomeric hydrolases exhibit a smaller mean activation free energy (16.8 vs 17.8 kcal/mol, 91.8% confidence level) than hetero-oligomeric hydrolases, they bind substrates less well (−5.6 vs −6.4 kcal/mol, 94.3% confidence level).

Dependence of the Enzyme Parameters on the Monomer Size. Even among monomeric hydrolases, the size varies significantly ranging from less than 200 residues to as large as 3000 residues, so are the larger/heavier enzymes more efficient than smaller ones? To evaluate how ΔG_{bind}, $\Delta G^{\neq}_{cat}$, and $\Delta G^{\neq}$ depend on the enzyme size, we computed their mean values and standard deviations for small (<275 aa), medium (275−500 aa), and large (500−3000) monomeric hydrolases. These three size ranges were chosen to ensure that each group contained at least 20 entries with roughly equal numbers (Table 5), enabling a fair comparison between small, medium, and large or very large enzymes. The results in Table 5 show that increasing the size of monomeric hydrolases does not seem to have a big impact on the catalytic efficiency. This is because increasing the number of aa residues not only lowered the activation free energy (smaller $\Delta G^{\neq}_{cat}$) but also attenuated substrate binding (ΔG_{bind} becomes less negative). The same trend was observed when the analysis was performed with the enzyme's molecular weight instead of the number of aa residues, or using different intervals in terms of the number of aa.

Table 5. Mean $\Delta G^{\neq}$, $\Delta G^{\neq}_{cat}$, and ΔG_{bind} for Monomeric Hydrolases as a Function of the Number of aa

no. of aa residues	no. of enzymes	average $\Delta G^{\neq}$ (kcal/mol)	average $\Delta G^{\neq}_{cat}$ (kcal/mol)	average ΔG_{bind} (kcal/mol)
0−275	27	10.8 ± 1.9	17.0 ± 2.1	−6.2 ± 1.4
275−500	31	10.2 ± 1.6	15.7 ± 1.8	−5.5 ± 2.8
500−3000	30	10.3 ± 1.8	15.5 ± 1.6	−5.2 ± 1.8

Dependence of the Enzyme Parameters on the Cellular Location. Hydrolase enzymes can operate within the cell or in the extracellular milieu. The availability of water in a cell, which contains many other proteins and molecular crowders, differs from that in the extracellular environment. As the availability of water is fundamental for hydrolysis reactions, we computed the mean $\Delta G^{\neq}$, $\Delta G^{\neq}_{cat}$, and ΔG_{bind}, and corresponding standard deviations for hydrolases acting inside and outside the cell. The results in Table 6 show that compared to intracellular hydrolases, extracellular hydrolases are more efficient (10.3 vs 11.5 kcal/mol, 99.5% confidence level) with significantly lower average activation free energy (15.6 vs 16.9 kcal/mol, 100% confidence level), but similar mean binding free energies (−5.4 vs −5.3 kcal/mol); the confidence level that

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887

Table 6. Mean $\Delta G^{\neq}$, $\Delta G^{\neq}_{cat}$, and ΔG_{bind} for Intracellular and Extracellular Hydrolases

cellular location	no. of enzymes	average $\Delta G^{\neq}$ (kcal/mol)	average $\Delta G^{\neq}_{cat}$ (kcal/mol)	average ΔG_{bind} (kcal/mol)
intracellular hydrolases	123	11.5 ± 2.2	16.9 ± 2.2	−5.4 ± 1.7
extracellular hydrolases	54	10.3 ± 2.4	15.6 ± 1.5	−5.3 ± 2.3
two-tailed *t*-test confidence level		99.5%	100.0%	21.1%

the mean binding free energies are statistically different is only 21%.

Dependence of the Enzyme Parameters on Environmental Conditions. To assess the influence of the environment on the activity of hydrolase enzymes, we have analyzed the influence of temperature as an energy source for reactions. Based on the organism corresponding to each entry, hydrolases were grouped into two classes: (1) Mesophilic enzymes with optimal activity at normal temperatures (20−45 °C) and (2) thermophilic enzymes with optimal activity at high or very high temperatures (>45 °C, typically over 60 °C). The latter class included a variety of hydrolase enzymes from thermophiles such as *Sulfolobus solfataricus*, *Pyrococcus furiosus*, *Thermus yunnanensis*, and *Alicyclobacillus acidocaldarius*. The results in Table 7 show that thermophilic hydrolases are significantly less

Table 7. Mean $\Delta G^{\neq}$, $\Delta G^{\neq}_{cat}$, and ΔG_{bind} for Mesophilic and Thermophilic Hydrolases

type	no. of enzymes	average $\Delta G^{\neq}$ (kcal/mol)	average $\Delta G^{\neq}$ (kcal/mol)	average ΔG_{bind} (kcal/mol)
mesophilic hydrolases	282	10.9 ± 2.4	16.5 ± 2.2	−5.7 ± 1.9
thermophilic hydrolases	30	13.2 ± 2.9	17.3 ± 2.2	−4.1 ± 2.5
two-tailed *t*-test confidence level		99.97%	91.3%	99.7%

efficient than mesophilic hydrolases (13.2 vs 10.9 kcal/mol, 99.97% confidence level). This is because compared to thermophilic hydrolases, mesophilic hydrolases bind substrates significantly better (−5.7 vs −4.1 kcal/mol, 99.7% confidence level) and have on average lower activation free energies (16.5 vs 17.3 kcal/mol, 91.3% confidence level).

Dependence of Enzyme Parameters on Substrate Specificity. Some hydrolases act on a single specific substrate, while others are promiscuous, catalyzing the hydrolysis of several compounds or even a whole class of molecular substrates (e.g., triacylglycerol lipase, E.C. 3.1.1.3, can act on a large number of different triacylglycerols). Do substrate-specific hydrolases bind substrate more tightly and/or have lower activation barriers than promiscuous hydrolases? To investigate the effect of substrate specificity of hydrolases on the ΔG_{bind}, $\Delta G^{\neq}_{cat}$, and $\Delta G^{\neq}$ values, hydrolases were divided into two groups: (1) those acting on a single specific substrate, and (2) those acting on more than one substrate or on a class of substrates. For each group, we computed the mean $\Delta G^{\neq}_{cat}$, ΔG_{bind}, and $\Delta G^{\neq}$ (Table 8). Not surprisingly, substrate-specific hydrolases with lower $\Delta G^{\neq}$ seem to be more catalytically competent than promiscuous ones (10.5 vs 11.2 kcal/mol, 86% confidence level), as they tend to bind substrate more tightly (−6.2 vs −5.4 kcal/mol, 98% confidence level) than promiscuous hydrolases. This trend is exemplified by the EC3.1 esterases: Substrate-specific esterases tend to not only bind substrate more tightly (−7.1 vs −6.1 kcal/mol, 99.2% confidence level), but also have on average lower $\Delta G^{\neq}_{cat}$ (16.0 vs 17.0 kcal/mol, 98.9% confidence level) and thus lower $\Delta G^{\neq}$ (8.9 vs 10.9 kcal/mol, 100% confidence level) than promiscuous esterases. The other subclasses contain insufficient (<20) substrate-specific or promiscuous hydrolases for a statistically sound comparison of their enzyme parameters.

Dependence of Enzyme Parameters on the Substrate Properties. Hydrolases use a water molecule to cleave a limited number of bond types in a large variety of substrate molecules. Even within the same subclass, the substrates can vary significantly in terms of their size, hydrogen bonding ability, flexibility, hydrophobicity, charge, and solvent accessibility. Do the different physical and structural properties of the substrates affect the catalytic efficiency and if so, is it via $\Delta G^{\neq}_{cat}$ or ΔG_{bind}? To answer this question, we correlated the mean $\Delta G^{\neq}$, $\Delta G^{\neq}_{cat}$, and ΔG_{bind} values for all hydrolases and the different subclass hydrolases with the substrate's molecular weight, volume, number of atoms, diameter, number of rotatable bonds, number of hydrogen-bond donors and acceptors, log *P*, log *S*, accessible surface area and corresponding decomposition into positive, negative, hydrophobic, and polar fractions (see Methods). The degree of correlation was expressed by Pearson's correlation coefficient, r^2. A value close to +1 indicates a positive correlation; that is, an increase in a given substrate property is associated with an increase in $\Delta G^{\neq}_{cat}$ or ΔG_{bind}, and thus a decrease in enzyme efficacy. A value close to −1 indicates a negative correlation; that is, an increase in a given substrate property is associated with lower $\Delta G^{\neq}_{cat}$ or more negative ΔG_{bind}, thus enhancing catalytic efficiency. A value of 0 suggests that the substrate property is independent of $\Delta G^{\neq}_{cat}$ or ΔG_{bind}. Although the mean ΔG_{bind}, $\Delta G^{\neq}_{cat}$, and $\Delta G^{\neq}$ for all hydrolases do not exhibit any strong correlations with the substrate properties evaluated ($|r^2| \leq 0.3$), those for the different subclass hydrolases do exhibit significant correlations with various substrate properties (Table 9). Below, we highlight those substrate properties that correlate with $\Delta G^{\neq}$, $\Delta G^{\neq}_{cat}$, or ΔG_{bind} for the different subclass hydrolases with $|r^2| \geq 0.5$.

Table 8. Average Values of $\Delta G^{\neq}_{cat}$, ΔG_{bind}, and $\Delta G^{\neq}$ for Substrate-Specific and Promiscuous Hydrolases

	no. of enzymes		mean $\Delta G^{\neq}$		mean $\Delta G^{\neq}_{cat}$		mean ΔG_{bind}	
no. of substrates	1	>1	1	>1	1	>1	1	>1
all	**60**	**279**	**10.5 ± 2.6**	**11.2 ± 2.5**	**16.7 ± 2.1**	**16.6 ± 2.2**	**−6.2 ± 1.8**	**−5.4 ± 2.0**
3.1 ester	19	101	8.9 ± 1.8	10.9 ± 2.5	16.0 ± 1.3	17.0 ± 2.5	−7.1 ± 1.2	−6.1 ± 1.9
3.2 sugars	2	65	8.1 ± 1.1	11.1 ± 1.8	13.2 ± 0.8	15.8 ± 1.5	−5.1 ± 0.3	−4.7 ± 1.9
3.4 peptide	8	51	12.0 ± 3.5	11.0 ± 2.2	17.4 ± 1.3	16.2 ± 1.4	−5.4 ± 3.3	−5.2 ± 1.9
3.5 C−N bonds	11	49	11.3 ± 2.1	12.1 ± 3.5	16.2 ± 1.8	16.7 ± 2.4	−4.9 ± 0.9	−4.6 ± 2.0
3.6 acid anhydrides	15	12	11.4 ± 2.5	11.5 ± 2.2	17.8 ± 2.9	17.8 ± 2.6	−6.5 ± 1.1	−6.2 ± 1.5

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887

Table 9. Pearson Correlation Coefficient Values of $\Delta G^{\neq}_{cat}$, ΔG_{bind}, and $\Delta G^{\neq}$ with Different Substrate Properties Calculated for Hydrolases in the Most Represented Subclasses[a]

Number of substrates	EC3.1 Ester			EC3.2 Sugars			EC3.4 Peptide			EC3.5 C–N Bonds			EC3.6 Acid Anhydrides		
	$\Delta G^{\neq}$	$\Delta G^{\neq}_{cat}$	ΔG_{bind}	$\Delta G^{\neq}$	$\Delta G^{\neq}_{cat}$	ΔG_{bind}	$\Delta G^{\neq}$	$\Delta G^{\neq}_{cat}$	ΔG_{bind}	$\Delta G^{\neq}$	$\Delta G^{\neq}_{cat}$	ΔG_{bind}	$\Delta G^{\neq}$	$\Delta G^{\neq}_{cat}$	ΔG_{bind}
Molecular Weight	0.7	0.8	0.0	0.0	−0.1	0.0	−0.4	0.1	−0.5	−0.6	−0.4	−0.6	−0.1	0.0	−0.1
Volume	0.7	0.8	0.0	0.0	−0.1	0.0	−0.4	0.1	−0.5	−0.6	−0.4	−0.6	−0.2	−0.2	0.0
# of Atoms	0.7	0.8	0.0	0.0	−0.1	0.0	−0.3	0.1	−0.5	−0.6	−0.4	−0.6	−0.2	−0.2	0.0
Diameter	0.7	0.8	−0.1	−0.1	−0.1	0.0	−0.3	0.1	−0.4	−0.5	−0.3	−0.6	−0.3	−0.3	−0.1
# of Rotatable Bonds	0.7	0.8	0.0	0.0	0.0	0.0	−0.3	0.1	−0.4	−0.4	−0.2	−0.5	−0.3	−0.2	−0.1
# of H-bond Acceptors	0.7	0.7	0.1	0.0	−0.1	0.1	−0.4	0.1	−0.5	−0.5	−0.4	−0.5	0.0	0.1	−0.2
# of H-bond Donors	0.6	0.7	−0.1	0.1	−0.1	0.2	−0.2	0.1	−0.3	−0.2	−0.2	−0.1	0.1	0.2	−0.2
ASA	0.7	0.8	0.0	−0.1	−0.1	0.0	−0.3	0.1	−0.4	−0.6	−0.3	−0.6	−0.3	−0.4	0.3
Fraction of ASA+	0.0	0.2	−0.3	−0.3	0.3	−0.5	0.3	0.2	0.2	0.3	0.2	0.2	0.2	0.1	0.1
Fraction of ASA–	0.0	0.0	0.0	0.0	−0.1	0.1	−0.4	0.0	−0.4	−0.1	0.0	−0.3	0.2	0.3	−0.2
Fraction of Polar ASA	0.2	0.3	−0.1	0.2	0.2	0.1	0.2	0.0	0.3	0.0	−0.1	0.1	0.5	0.6	−0.2
Fraction of Hydrophobic ASA	−0.2	−0.3	0.1	−0.2	−0.2	−0.1	−0.2	0.0	−0.3	0.0	0.1	−0.1	−0.5	−0.6	0.2
Log S	−0.3	−0.4	0.1	0.3	−0.1	0.4	0.5	−0.1	0.6	0.3	0.1	0.4	0.0	−0.1	0.2
Log P	−0.4	−0.4	−0.1	−0.1	0.1	−0.2	−0.1	0.0	−0.1	0.1	0.2	0.0	−0.3	−0.2	−0.2

[a]Positive and negative correlation coefficients with magnitude ≥0.5 are in blue and red, respectively.

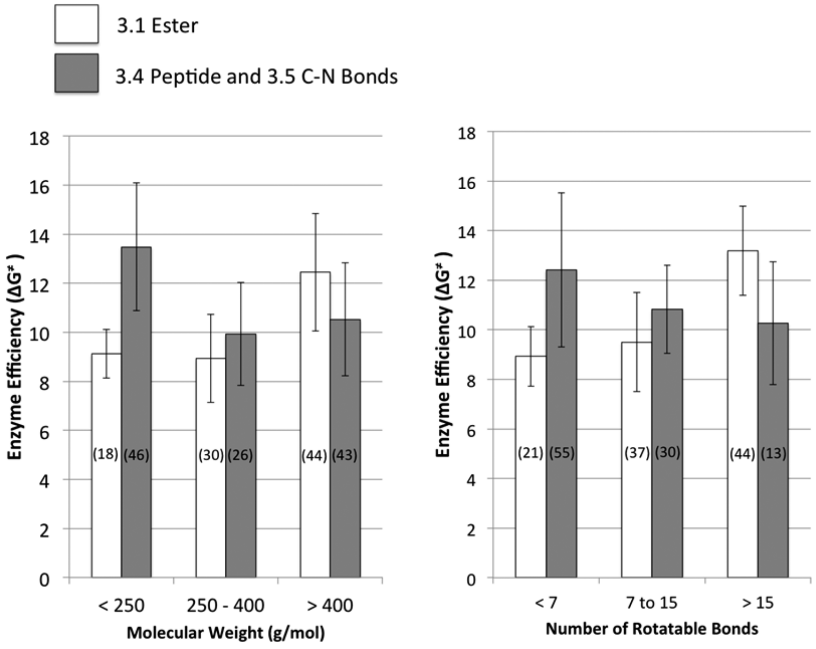


Figure 4. Variation of the mean $\Delta G^{\neq}$ for EC3.1 (white bars) and EC3.4/EC3.5 (gray bars) hydrolases with the molecular weight (left) and number of rotatable bonds (right) of the substrates. The numbers in parentheses denote the number of $\Delta G^{\neq}$ values corresponding to the given range of molecular weight or number of rotatable bonds.

Substrate Size, Flexibility, and Hydrogen-Bonding Propensity. Larger substrates with more degrees of freedom tend to be more flexible. Substrate flexibility could be detrimental for catalysis, as it might result in greater loss of conformational entropy upon binding to an enzyme and thus less favorable binding free energy, ΔG_{bind}. However, no apparent correlation between substrate size/flexibility and higher binding free energies was observed for all hydrolases. On the contrary, increasing substrate size (molecular weight, volume, number of atoms, diameter) flexibility (number of rotatable bonds) correlated with tighter binding by EC3.4 and EC3.5 hydrolases that act on C–N bonds (r^2 = −0.5 or −0.6), and thus more proficient enzymes. Substrate size/flexibility could nevertheless be detrimental for catalysis by yielding higher activation free energies for certain enzymes. Indeed, increasing the substrate size and flexibility correlated with higher activation free energies for EC3.1 esterases (r^2 = 0.8), and thus less competent enzymes.

That increasing substrate size/flexibility generally yielded less competent EC3.1 esterases but more proficient EC3.4/EC3.5 enzymes is evident in Figure 4, which shows the variation of the average enzyme efficiency values for EC3.1 vs EC3.4/EC3.5

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887

hydrolases as a function of the molecular weight and number of rotatable bonds of the corresponding substrates. The data for EC3.4/EC3.5 hydrolases were combined, as separating them resulted in <20 $\Delta G^{\neq}$ values for certain molecular weight or flexibility range; for example, none of the EC3.5 hydrolase substrates had over 15 rotatable bonds. Figure 4 shows that EC3.1 hydrolases acting on large esters (>400 g/mol) have an average $\Delta G^{\neq}$ (~12.5 ± 2.4 kcal/mol) that is 33% larger than those acting on smaller esters (<400 g/mol), whose mean $\Delta G^{\neq}$ is ~9 ± 1.5 kcal/mol. A similar, but even more pronounced trend can be observed when the mean $\Delta G^{\neq}$ for EC3.1 esterases are plotted as a function of the number of rotatable bonds: esterases acting on flexible substrates with >15 rotatable bonds exhibit a considerably higher $\Delta G^{\neq}$ value (13.2 ± 1.8 kcal/mol) than those with <15 rotatable bonds (~9 ± 2 kcal/mol). In sharp contrast, EC3.4/EC3.5 enzymes show significantly larger mean $\Delta G^{\neq}$ values for small substrates with molecular weight <250 g/mol and <7 rotatable bonds, as compared to larger substrates.

An increasing substrate's ability to establish hydrogen bonds and to be solvent accessible also yielded less competent EC3.1 esterases, but more proficient EC3.5 enzymes: Increasing the substrate's number of hydrogen-bond acceptors and accessible surface area correlated with higher activation free energies for EC3.1 esterases ($r^2 \geq 0.7$), but tighter binding by EC3.5 hydrolases ($r^2 = -0.5$ or -0.6). These trends are also found for EC3.1/EC3.5 hydrolases that function without cofactors.

What factors could account for the different trends observed for EC.3.1 and EC3.4/EC3.5 hydrolases? One plausible explanation why EC3.1 esterases become less competent with increasing substrate size/flexibility is that large lipidic substrates need to undergo a preorganization before reacting, as they are transferred from an environment where they are typically stably packed in micelles to one where they associate with the enzyme with an energetic cost. In contrast to EC3.1 esterases, EC3.4/EC3.5 hydrolases become more proficient with increasing substrate size/flexibility probably because large peptide/protein substrates might be preorganized by folding into particular conformations, thus restricting the number of freely rotatable bonds and stabilizing the reactant state of the enzyme.

Substrate Charge, Polarity, and Solubility. The fraction of positively or negatively charged, accessible surface area of the substrate does not seem to affect catalytic efficiency. This suggests that differences in the substrate's positive/negative charge can be buffered by the respective enzyme structures. On the other hand, increasing the fraction of hydrophobic accessible surface area enhanced catalytic efficiency of EC3.6 hydrolases that act on acid anhydrides such as diphosphatases, triphosphatases, and GTPases, as acid anhydrides with a higher fraction of hydrophobic ASA ($r^2 = -0.6$) and a smaller fraction of polar ASA ($r^2 = 0.6$) resulted in lower activation free energies. Solubility seems to affect the catalytic efficiency of EC3.4 peptidases: less soluble substrates bind appreciably better, indicating the importance of desolvation in substrate binding for this class of enzyme.

DISCUSSION

We have presented a detailed analysis of the kinetic, structural, and biological experimental information currently available for hydrolases in an attempt to reveal specific trends, similarities and differences in the catalytic activity of this large and quite diverse class of enzymes. Here, we discuss the key findings and highlight questions that remain unaddressed.

Narrowing ΔG_{bind} and $\Delta G^{\neq}_{cat}$. Considering the huge diversity of substrate and enzyme structures, the binding free energies ΔG_{bind} of hydrolases fall in a limited range with a mean of −5.5 ± 2.0 kcal/mol for 339 enzymes. The latter is far from the maximum binding free energy observed for all enzymes, −18 kcal/mol,[22] indicating that hydrolases are mediocre binders. This is probably because binding the substrate too tightly might delay product release. The activation free energies $\Delta G^{\neq}_{cat}$ of hydrolases also fall in a narrow range with a mean of 16.6 ± 2.2 kcal/mol for 339 enzymes. Hence, hydrolases do more than lowering the energy barriers for catalysis: they lower the apparent activation free energies $\Delta G^{\neq}_{cat}$ for different chemical reactions by different degrees to a narrow range, as compared to the corresponding activation free energy range in solution. This finding, based on 339 hydrolase reactions with verified natural substrates, confirms an earlier observation based on only 11 hydrolases,[16] where the k_{cat} values range from 3.8 to 10^6 s^{-1}, but the rates of the corresponding uncatalyzed reactions (10^{-20} to 10^{-1} s^{-1}) span ~19 orders of magnitude. Due to the limited range in $\Delta G^{\neq}_{cat}$ and ΔG_{bind} as well as to some correlation/compensation between them, the catalytic efficiency is also confined to a very narrow range.

Different Subclasses, Similar Enzyme Efficiency. Different hydrolase subclasses use different strategies to achieve comparable catalytic efficiency by lowering the activation free energy or the binding free energy to varying extents: the mean $\Delta G^{\neq}_{cat}$ and ΔG_{bind} vary more from subclass to subclass than the catalytic efficiency, $\Delta G^{\neq}$ (Table 1). In general, esterases (EC3.1) as well as hydrolases acting on nonpeptidic C−N bonds (EC3.5) and acid anhydrides (EC3.6) tend to improve catalytic efficiency via lowering of the activation free energy rather than tighter binding of their substrates. Conversely, glycosylases (EC3.2) and proteases/peptidases (EC3.4) tend to enhance catalytic efficiency by binding the substrate more tightly (Table 2). The different dependencies of $\Delta G^{\neq}_{cat}$, ΔG_{bind}, and $\Delta G^{\neq}$ with the different substrate properties also indicate how enzymes adopt a variety of different strategies to bind different substrates and to hydrolyze efficiently different chemical reactions, resulting in rather uniform values of $\Delta G^{\neq}_{cat}$, ΔG_{bind}, and $\Delta G^{\neq}$.

Improving Substrate Binding ΔG_{bind}. How can substrate binding by hydrolases be improved? One obvious way is for hydrolases to bind a single specific substrate rather than multiple substrates (Table 8). Another way is to employ Mg^{2+} as a cofactor (Table 3). The substrate itself could also contribute to its tighter binding. Large, insoluble substrates bind better to EC3.4 and EC3.5 hydrolases that hydrolyze C−N bonds than small, soluble ones.

Lowering the Activation Barrier, $\Delta G^{\neq}_{cat}$. How can the activation barrier be lowered? This can be achieved using a bimetallic center such as a binuclear Zn-site to stabilize the rate-limiting transition-state structure (Table 3) and to employ single chain enzymes rather than hetero-oligomeric ones. The size of the enzyme or substrate could also contribute to a lowering of the activation barrier. Activation free energies tend to be lower for large (>275 aa residues) monomeric enzymes as opposed to smaller ones (Table 5). They also tend to be lower for EC3.1 esterases that hydrolyze small, rigid, and buried substrates rather than large, flexible, solvent-accessible ones (Table 9).

Enhancing Catalytic Efficiency, $\Delta G^{\neq}$. Clearly, factors that improve substrate binding and/or lower the activation barrier would enhance catalytic efficiency. Thus, the presence of

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887

binuclear Zn-sites or mononuclear Mg^{2+} or Ca^{2+} sites tends to improve the catalytic efficiency of hydrolases (Figure 3). Large monomeric hydrolases with a mean $\Delta G^{\neq}$ = 10.3 kcal/mol (Table 5) appear to be more efficient than increasing enzyme size via oligomerization (~11.5 kcal/mol, Table 4). Substrate-specific hydrolases seem more competent than promiscuous ones that have to act on more than one substrate (Table 8). The mean $\Delta G^{\neq}$ values for the subclasses in Table 1 appear more similar than the activation or binding free energies, as they were not different within a confidence level of 95% except for the comparison between the mean $\Delta G^{\neq}$ values of EC3.1 (10.6 kcal/mol) and EC3.5 (11.9 kcal/mol) hydrolases.

Limitations and Future Work. The sheer size and diversity of data and sources in the BRENDA database would likely yield some incorrect substrate or kinetic data. Hence, in this work, all the entries were cross-checked against other databases and manually inspected to minimize errors. Although laborious and time-consuming, such a process is necessary, but it still does not guarantee that the final data set used is 100% error-proof. Hence, it is critical to employ appropriate statistical tests to identify statistically significant findings and limit the chances that the key conclusions would change due to errors in the data set. We are currently evaluating other classes of enzymes to verify which of the trends found for hydrolases might be generic or unique to a given enzyme class.

CONCLUSIONS

Hydrolases are involved in breaking different chemical bonds of various biological substrates, including proteins, carbohydrates, lipids, and nucleic acids of very different sizes and complexity. Despite the large number/diversity of chemical reactions, enzymes, and substrates, hydrolases have characteristic activation free energies, substrate binding free energies, and catalytic efficiencies.

Different subclasses of hydrolases employ different strategies to achieve these values, including the use of cofactors, different numbers of chains, different chains, and different numbers of aa residues. The large structural and physicochemical diversity of the substrates upon which they act is largely neutralized by the different hydrolases, resulting in activation free energies, binding free energies, and enzyme efficiencies that are in general almost independent (Table 10).

Table 10. Average Values of Hydrolases for Enzymatic Properties and Main Conclusions

characteristic properties	average data for hydrolases
$\Delta G^{\neq}_{cat}$ (kcal/mol)	16.6 ± 2.2
ΔG_{bind} (kcal/mol)	−5.5 ± 2.0
subclasses	even though different subclasses exhibit different dependencies of $\Delta G^{\neq}_{cat}$, ΔG_{bind} or $\Delta G^{\neq}$ on the substrate properties, their overall outcome is very similar.
cofactors	hydrolases with cofactors have lower activation free energies and lower binding free energies, but on average by only 0.6 kcal/mol.
size	smaller enzymes (<250 aa residues) are efficient in binding substrates but not in lowering the activation free energy.
oligomerization	monomeric hydrolases are more catalytic efficient with lower activation free energies and binding free energies than oligomeric hydrolases.
substrate specificity	substrate-specific hydrolases bind substrates better with an average binding free energy of 0.8 kcal/mol lower than promiscuous hydrolases.

We might think that human metabolism is systemic and all biochemical transformations are interconnected and interdependent. Interestingly, the other nonhuman species analyzed herein also work in the same biochemical time scale as humans, probably because the chemistry of life as a whole is interdependent, within an organism and across all organisms. The chemistry of life seems to have its own clock, its own rhythm, which is precise, narrow, and absolutely synchronized. Enzymes can hence be pictured as antidemocratic tyrants, masters in neutralizing differences and forcing an energetic convergence to assume total control of the reaction they catalyze or as perfect machines capable of integrating the plethora of reactions involving all sorts of substrates with precision in an interdependent world.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: carmay@gate.sinica.edu.tw.

*E-mail: pafernan@fc.up.pt.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work has been funded by FEDER/COMPETE and Fundação para a Ciência e a Tecnologia through programs PEst-C/EQB/LA0006/2011, EXCL/QEQ-COM/0394/2012, IF/00052/2014.

REFERENCES

(1) (a) Zhang, X. Y.; Houk, K. N. *Acc. Chem. Res.* **2005**, *38*, 379−385. (b) Hammes, G. G.; Benkovic, S. J.; Hammes-Schiffer, S. *Biochemistry* **2011**, *50*, 10422−10430.

(2) (a) Bruice, T. C.; Benkovic, S. J. *Biochemistry* **2000**, *39*, 6267−6274. (b) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H. B.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210−3235. (c) Olsson, M. H. M.; Parson, W. W.; Warshel, A. *Chem. Rev.* **2006**, *106*, 1737−1756. (d) Villa, J.; Warshel, A. *J. Phys. Chem. B* **2001**, *105*, 7887−7907. (e) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186−195. (f) Benkovic, S. J.; Hammes-Schiffer, S. *Science* **2003**, *301*, 1196−1202. (g) Kohen, A.; Klinman, J. P. *Acc. Chem. Res.* **1998**, *31*, 397−404. (h) Nagel, Z. D.; Klinman, J. P. *Nat. Chem. Biol.* **2009**, *5*, 543−550. (i) Marti, S.; Roca, M.; Andres, J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J. *Chem. Soc. Rev.* **2004**, *33*, 98−107.

(3) Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. U. S. A.* **1971**, *68*, 1678−1683.

(4) Menger, F. M.; Glass, L. E. *J. Am. Chem. Soc.* **1980**, *102*, 5404−5406.

(5) Bruice, T. C. *Acc. Chem. Res.* **2002**, *35*, 139−148.

(6) Dewar, M. J. S.; Storch, D. M. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 2225−2229.

(7) Cleland, W. W.; Kreevoy, M. M. *Science* **1994**, *264*, 1887−1890.

(8) Agarwal, P. K. *J. Am. Chem. Soc.* **2005**, *127*, 15248−15256.

(9) Pauling, L. *Nature* **1948**, *161*, 707−709.

(10) (a) Blow, D. *Structure* **2000**, *8*, R77−R81. (b) Ringe, D.; Petsko, G. A. *Science* **2008**, *320*, 1428−1429. (c) Williams, D. H.; Stephens, E.; Zhou, M. *Chem. Commun.* **2003**, 1973−1976. (d) Hammes, G. G. *J. Biol. Chem.* **2008**, *283*, 22337−22346. (e) Schowen, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 11931−11932. (f) Giraldo, J.; Roche, D.; Rovira, X.; Serra, J. *FEBS Lett.* **2006**, *580*, 2170−2177. (g) Graham, J. D.; Buytendyk, A. M.; Wang, D.; Bowen, K. H.; Collins, K. D. *Biochemistry* **2014**, *53*, 344−349.

(11) (a) Zinovjev, K.; Ruiz-Pernia, J. J.; Tunon, I. *J. Chem. Theory Comput.* **2013**, *9*, 3740−3749. (b) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12431−12441. (c) Ramos, M. J.; Fernandes, P. A. *Acc. Chem. Res.* **2008**, *41*, 689−698. (d) Kamerlin, S. C. L.; Haranczyk, M.; Warshel, A. *J. Phys. Chem. B*

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887

2009, *113*, 1253−1272. (e) Field, M. J. *J. Comput. Chem.* **2002**, *23*, 48−58. (f) Mulholland, A. J. *Drug Discovery Today* **2005**, *10*, 1393−1402. (g) Himo, F. *Theor. Chem. Acc.* **2006**, *116*, 232−240. (h) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425−443. (i) van der Kamp, M. W.; Mulholland, A. J. *Nat. Prod. Rep.* **2008**, *25*, 1001−1014. (j) Leopoldini, M.; Marino, T.; Michelini, M. D.; Rivalta, I.; Russo, N.; Sicilia, E.; Toscano, M. *Theor. Chem. Acc.* **2007**, *117*, 765−779. (k) Ranaghan, K. E.; Mulholland, A. J. *Int. Rev. Phys. Chem.* **2010**, *29*, 65−133. (l) Gerlt, J. A.; Allen, K. N.; Almo, S. C.; Armstrong, R. N.; Babbitt, P. C.; Cronan, J. E.; Dunaway-Mariano, D.; Imker, H. J.; Jacobson, M. P.; Minor, W.; Poulter, C. D.; Raushel, F. M.; Sali, A.; Shoichet, B. K.; Sweedler, J. V. *Biochemistry* **2011**, *50*, 9950−9962. (m) Markwick, P. R. L.; McCammon, J. A. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20053−20065.

(12) McDonald, A. G.; Tipton, K. F. *FEBS J.* **2014**, *281*, 583−592.

(13) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(14) (a) Vuong, T. V.; Wilson, D. B. *Biotechnol. Bioeng.* **2010**, *107*, 195−205. (b) Trincone, A.; Giordano, A. *Curr. Org. Chem.* **2006**, *10*, 1163−1193.

(15) Siirola, E.; Frank, A.; Grogan, G.; Kroutil, W. *Adv. Synth. Catal.* **2013**, *355*, 1677−1691.

(16) (a) Wolfenden, R. *Chem. Rev.* **2006**, *106*, 3379−3396. (b) Wolfenden, R. *Annu. Rev. Biochem.* **2011**, *80*, 645−667.

(17) (a) Schomburg, I.; Chang, A.; Placzek, S.; Sohngen, C.; Rother, M.; Lang, M.; Munaretto, C.; Ulas, S.; Stelzer, M.; Grote, A.; Scheer, M.; Schomburg, D. *Nucleic Acids Res.* **2013**, *41*, D764−D772. (b) Schomburg, I.; Chang, A.; Schomburg, D. *Nucleic Acids Res.* **2002**, *30*, 47−49.

(18) (a) Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. *Nucleic Acids Res.* **2003**, *31*, 3784−3788. (b) Artimo, P.; Jonnalagedda, M.; Arnold, K.; Baratin, D.; Csardi, G.; de Castro, E.; Duvaud, S.; Flegel, V.; Fortier, A.; Gasteiger, E.; Grosdidier, A.; Hernandez, C.; Ioannidis, V.; Kuznetsov, D.; Liechti, R.; Moretti, S.; Mostaguir, K.; Redaschi, N.; Rossier, G.; Xenarios, I.; Stockinger, H. *Nucleic Acids Res.* **2012**, *40*, W597−W603.

(19) Warshel, A. *J. Biol. Chem.* **1998**, *273*, 27035−27038.

(20) Bairoch, A.; Boeckmann, B. *Nucleic Acids Res.* **1991**, *19*, 2247−2248.

(21) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. *J. Cheminf.* **2011**, *3*, 1−14.

(22) Smith, R. D.; Engdahl, A. L.; Dunbar, J. B.; Carlson, H. A. *J. Chem. Inf. Model.* **2012**, *52*, 2098−2106.

(23) Hu, L. G.; Benson, M. L.; Smith, R. D.; Lerner, M. G.; Carlson, H. A. *Proteins: Struct., Funct., Genet.* **2005**, *60*, 333−340.

(24) (a) Dudev, T.; Lim, C. *Annu. Rev. Biophys.* **2008**, *37*, 97−116. (b) Dudev, T.; Lim, C. *J. Phys. Chem. B* **2009**, *113*, 11754−11764. (c) Sousa, S. F.; Lopes, A. B.; Fernandes, P. A.; Ramos, M. J. *Dalton Trans.* **2009**, 7946−7956. (d) Dudev, T.; Lim, C. *Chem. Rev.* **2014**, *114*, 538−556.

(25) Yang, T. Y.; Dudev, T.; Lim, C. *J. Am. Chem. Soc.* **2008**, *130*, 3844−3852.

DOI: 10.1021/acscatal.5b00923
ACS Catal. 2015, 5, 5877−5887