

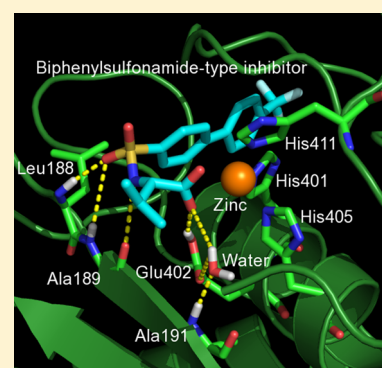
Combined QM/MM (ONIOM) and QSAR Approach to the Study of Complex Formation of Matrix Metalloproteinase-9 with a Series of Biphenylsulfonamides—LERE-QSAR Analysis (V)

Tatsusada Yoshida, Seiji Hitaoka, Akira Mashima, Takuya Sugimoto, Hiroshi Matoba, and Hiroshi Chuman*

Institute of Health Biosciences, The University of Tokushima Graduate School, 1-78 Shomachi, Tokushima 770-8505, Japan

S Supporting Information

ABSTRACT: We previously proposed a novel QSAR (quantitative structure–activity relationship) procedure called LERE (linear expression by representative energy terms)-QSAR involving molecular calculations such as an ab initio fragment molecular orbital ones. In the present work, we applied LERE-QSAR to complex formation of matrix metalloproteinase-9 (MMP-9) with a series of substituted biphenylsulfonamides. The results shows that the overall free-energy change accompanying complex formation is due to predominantly the contribution from the electrostatic interaction with the zinc atom in the active site of MMP-9. Carbonic anhydrase (CA) belongs to the zinc-containing protease family. In contrast to the current case of MMP-9, the overall free-energy change during complex formation of CA with a series of benzenesulfonamides is due to the contributions from the solvation and dissociation free-energy changes, as previously reported. Comparison of the two sets of results indicates quantitative differences in the relative contributions of free-energy components to the overall free-energy change between the two data sets, corresponding with those in the respective classical QSAR equations. The LERE-QSAR procedure was demonstrated to quantitatively reveal differences in the binding mechanisms between the two cases involving similar but different zinc-containing proteins at the electronic and atomic levels.



1. INTRODUCTION

Zinc atoms are known to play an essential role in zinc-containing proteins. Up to 10% of the human genome encodes zinc-containing proteins, amounting to at least 3000 proteins,¹ and zinc-containing proteins comprise the most abundant class among metalloproteins. The zinc-containing proteins further comprise a number of families that exhibit varying degrees of structural similarity. Among a variety of zinc-containing proteins, thermolysin, carboxypeptidase A, carbonic anhydrase, angiotensin-converting enzyme, and TNF- α -converting enzyme have been well studied from the structural and catalytic mechanism points of view.^{2–8} There has been an increasing interest in zinc-containing proteins relating to a variety of human diseases, such as cancer, rheumatoid arthritis, and diabetes.⁹ For drug design targeting zinc-containing proteins, it is now desired and important to physicochemically and quantitatively understand the contribution of the zinc atom to the binding of drug-candidate compounds with zinc-containing proteins.

We previously reported correlation analyses of the binding affinity of a series of substituted benzenesulfonamides (BSAs) with carbonic anhydrase (CA), one of the zinc-containing proteins that catalyze the rapid interconversion of carbon dioxide and water to bicarbonate and protons, using ab initio FMO (fragment molecular orbital) and continuum solvation model calculations.^{10,11} We demonstrated that correlation analysis based on LERE (linear expression by representative energy

terms) formulation^{10–13} can excellently reproduce the observed variation of the free-energy change associated with the complex formation of BSAs and CA. The results revealed that changes in the intrinsic binding interaction, solvation, and dissociation free-energies determine the observed overall free-energy change.

Matrix metalloproteinases (MMPs) are also zinc-containing proteins. MMPs are responsible for normal turnover and remodeling of the extracellular matrix, and are capable of degrading most components in the extracellular matrix, such as collagen, laminin, fibronectin, elastin, and serpin.^{14,15} Currently, at least 26 human MMPs are known on the basis of their substrate specificity. MMP-9 (gelatinase B) is known as the significant contributor into several cardiovascular diseases, including atherosclerosis, hypertension, heart failure and ischemic heart disease.^{16,17}

Verma et al.¹⁸ reported a classical QSAR (quantitative structure–activity relationship) study as to the inhibitory potency of a series of biphenylsulfonamide derivatives (hereafter denoted as BSs) toward MMP-9. The QSAR equation according to their report is expressed by $\log(1/IC_{50}) = -1.14 \sigma + 1.24 MR + 4.38$ ($r = 0.944$), where σ and MR are the Hammett σ constant and molar refractivity of the substituent at the *para* position in

Received: June 4, 2012

Revised: July 20, 2012

Published: July 30, 2012

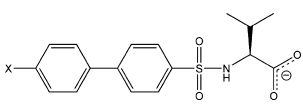
the biphenyl part, respectively. However, it requires a more direct and straight approach involving molecular calculations such as an ab initio MO ones on complexes of MMP-9 with BSs, in order to understand the atomic and electronic binding mechanisms.

In the present work, we examine the atomic and electronic mechanisms underlying binding between MMP-9 and BSs, using the LERE-QSAR procedure. The current results will be useful for the development of new drugs for disorders involving MMP-9.

2. METHODS

2.1. Compound Set. The compounds were basically the same as those in the paper by Verma et al.,¹⁸ but the *meta*- and *ortho*-substituted compounds were excluded because of the multiple conformations of their biphenyl rings. Table 1 lists the

Table 1. Chemical Structure of BSs with pIC_{50} , ΔG_{obs} , σ , and MR



compound		$pIC_{50}^{a,b}$	ΔG_{obs}^c	σ^d	MR ^d
no.	-X				
1	OMe	5.66	-7.72	-0.27	0.79
2	Me	5.41	-7.38	-0.17	0.56
3	Br	5.10	-6.95	0.23	0.89
4	CHO	4.77	-6.50	0.42	0.69
5	CF ₃	4.70	-6.41	0.54	0.50
6	H	4.59	-6.26	0.00	0.10
7	NO ₂	4.42	-6.03	0.78	0.74
8	CN	4.23	-5.77	0.66	0.63
9	F	4.19	-5.71	0.06	0.09

^a $pIC_{50} = -\log IC_{50}$ (in M). ^bTaken from ref 19. ^c $\Delta G_{obs} = 2.303 RT \log IC_{50}$ ($T = 298$ K), in kcal/mol. ^dTaken from ref 18.

chemical structures of the series of nine *para*-substituted BSs (biphenylsulfonamides) examined in this study, together with pIC_{50} ($= -\log IC_{50}$), overall free-energy change ΔG_{obs} ($= 2.303RT \log IC_{50}$, $T = 298$ K), Hammett σ , and molar reflectivity MR.

2.2. Complex Structure. The initial complex structure of the catalytic domain of MMP-9 with *para*-trifluoro substituted compound 5 ($X = CF_3$) was constructed based on the crystallographic structure of MMP-9 with *N*-[(4'-iodobiphenyl-4-yl) sulfonyl]-D-tryptophan having biphenyl, sulfonamide, and carboxylate groups shared with compound 5 (PDB code: 2OW0²⁰). Compound 5 was placed in MMP-9 so as to resemble the interactions of the ligand with residues and zinc ions in the crystallographic structure, as shown in Figure 1. The mutated amino acid residue, Gln402, in the active site was replaced by Glu402 in the wild-type MMP-9. Based on the pH dependence of MMP-3's catalytic and inhibitory activities, Johnson et al.²¹ deduced that a carboxylate type inhibitor binds to the protonated form of Glu202 (corresponding to Glu402 in MMP-9), although the carboxylate in a ligand takes on an anionized form. We further confirmed that the calculated pK_a value of Glu402 was 7.70 (at pH 7.0) in the complex of a carboxylate type inhibitor with wild-type MMP-9 (2OW0), by using the PDB2PQR server.²² These two results strongly indicate that Glu402 takes on a protonated form during complex formation. Therefore, we adopted the protonation form of Glu402 in the current study. The protonated states of Arg, Lys, Asp, and Glu residues having ionizable side

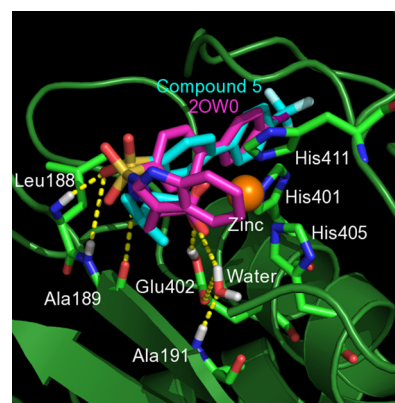


Figure 1. Superposition of compound 5 ($X = CF_3$; cyan) on *N*-[(4'-iodobiphenyl-4-yl) sulfonyl]-D-tryptophan (PDB code: 2OW0; magenta) in the active site of MMP-9.

chains other than Glu402 were treated as charged entities, according to their calculated pK_a values. We also examined all possible protonated and tautomeric states of three His residues (His401, His405, and His411) coordinating the zinc atom in the active site, using the PDB2PQR server, and then we adopted the N(δ)-H (neutral) form as the most probable state of the three His residues. Four calcium ions (Ca^{2+}), another zinc ion (the second zinc ion) located apart from the active site, and 118 water molecules found in 2OW0 were retained in the complex structure of MMP-9 with compound 5 (the calcium ions and the second zinc ion are known to function in maintaining the structure of MMPs²³). One water molecule crystallographically observed within the active site forms a hydrogen-bonding network with Glu402, Ala191, and the carboxylate in the ligand, suggesting the water molecule plays an important role in the catalytic reaction.

We performed two step refinement of the complex structure geometry. First, we optimized the complex structure of MMP-9 with compound 5 having the largest substituent CF_3 among compounds 1–9, using QM/MM (ONIOM in Gaussian 09 program²⁴) and MD calculations (AMBER11 package²⁵). A total 1.7 ns of MD calculations were carried out with atomic positional restraints on the two zinc binding sites, and then the average structure of the complex was generated from the last 1.0 ns trajectory. Other complex structures of MMP-9 with compounds 1–4 and 6–9, i.e., other than 5, were constructed by replacing the trifluoro group in compound 5 with the corresponding substituents.

Finally, the two-layer ONIOM model^{26–37} with the mechanical embedding (ME) scheme was applied for further geometry refinement of the complex structures of MMP-9 with compounds 1–9. The QM region includes the zinc ion located in the active site, an inhibitor, the water molecule in the active site, the side chain atoms of Leu397, Val398, His401, Glu402, His405, His411, Leu418, and Tyr423, and the main chain atoms of Leu188 and Ala189 in the active site, as illustrated in Figure 2. The HF/6-31G(d) level calculation was used to describe the QM part of the system, and the parm99 force field³⁸ and general AMBER force field (GAFF)³⁹ were used to describe the MM part of the system. The partial atomic charges for the zinc-coordinating environment in the active site were determined using a small model complex ($[Zn^{2+}(\text{methylimidazole})_3(\text{inhibitor}^-)]^+$), using the RESP (restrained electrostatic potential) fitting procedure⁴⁰ with HF/6-31G(d) calculations. The structure in the QM part of the complex was geometry

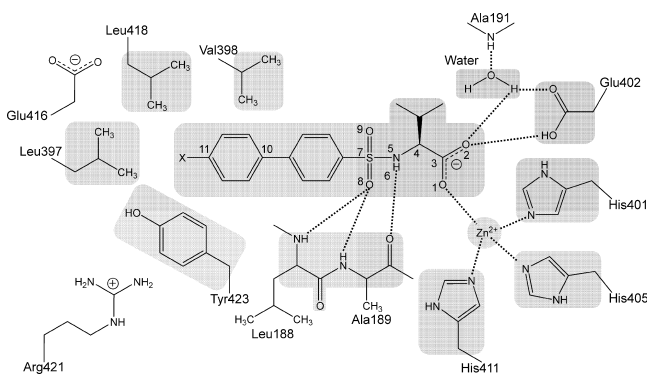


Figure 2. Schematic representation of the interaction of BS with amino acid residues, zinc ion, and water molecule in the active site of MMP-9. Atoms in shadow areas are treated as QM region in the ONIOM calculation.

optimized, whereas atoms in the MM part were fixed. The complex structures thus obtained were used for the calculation of LERE terms.

Table 2 lists interatomic distances among major atoms in the complex of compound **6** ($X = H$) with MMP-9, as well as the

Table 2. Interatomic Distance and RESP Charge in the Complex of Compound **6** ($X = H$) with MMP-9

atom A	RESP charge ^a	atom B	distance (Å) ^b
compound 6 -O (1) ^c	−0.4586	Zn	1.931
compound 6 -O (2) ^c	−0.4586	Glu402-H _e	1.754
		water-H	2.177
compound 6 -C (3) ^c	0.3822		
compound 6 -C (4) ^c	−0.1953		
compound 6 -N (5) ^c	−0.5560		
compound 6 -H (6) ^c	0.3522	Ala189-C=O	1.792
compound 6 -S (7) ^c	0.8123		
compound 6 -O (8) ^c	−0.4348	Leu188-NH	1.952
		Ala189-NH	2.426
compound 6 -O (9) ^c	−0.4348		
compound 6 -C (10) ^c	0.1222	Tyr423-C _β	4.187
		Val398-C _γ	5.852
compound 6 -C (11) ^c	−0.1757	Leu418-C _δ	4.136
		Leu397-C _β	5.547
		Arg421-C _β	6.238
		Glu416-C _β	10.048
Zn	0.8720		
His401-N _e	−0.3758	Zn	1.991
His405-N _e	−0.3758	Zn	2.009
His411-N _e	−0.3758	Zn	2.002
Ala191-NH	^d	water-O	1.973
Glu402-O _e	^d	water-H	2.800

^aIn esu. ^bIn Å. ^cA number in parentheses is the atomic number marked in Figure 2. ^dThe RESP charges in the TIP3P model⁴¹ and Amber parm99 force field were used for water molecules and amino acid residues other than the three His residues (401, 405, and 411), respectively.

RESP charges on major atoms. Tables S1-1–S1-9 in the Supporting Information list those of all of the compounds **1**–**9**. Also, Table S2 lists all the atomic coordinates in the full complex structure of compound **6** with MMP-9, together with the RESP charges.

2.3. LERE-QSAR Equation. The first assumption made in formulating the LERE-QSAR equation^{10–13} is that the free-energy terms comprising the overall free-energy change associated with complex formation are all additive.^{42,43}

$$\Delta G_{\text{obs}} = \Delta G_{\text{bind}} + \Delta G_{\text{sol}} + \Delta G_{\text{diss}} + \Delta G_{\text{others}}$$

(additivity assumption)

ΔG_{obs} on the left-hand side is the overall free-energy change calculated from the observed inhibitory potency. ΔG_{bind} , ΔG_{sol} , and ΔG_{diss} (typically taken as dominant free-energetic terms) are the intrinsic binding interaction energy of an inhibitor with a protein, the solvation free-energy change associated with complex formation, and the dissociation free-energy change of an ionizable inhibitor, respectively. ΔG_{others} , the sum of free-energy terms other than representative free-energy terms, ΔG_{bind} , ΔG_{sol} , and ΔG_{diss} , is assumed to be linear with that of representative free-energy terms (LERE approximation). ΔG_{others} consists of, for instance, the structural deformation (adaptation/induced fitting) energy associated with complex formation

$$\Delta G_{\text{others}} = \beta(\Delta G_{\text{bind}} + \Delta G_{\text{sol}} + \Delta G_{\text{diss}}) + \text{const (LERE assumption)}$$

where ΔG_{others} is expected to work as a penalty term ($\beta < 0$ and/or const > 0).

When treating a congeneric series of ligands, (1) the entropic contribution arising from conformational fluctuation to ΔG_{bind} is small,⁴⁴ (2) ΔG_{sol} and ΔG_{diss} , calculated with continuum solvation models, are replaced by their dominant polar (electrostatic) contributions $\Delta G_{\text{sol}}(\text{polar})$ and $\Delta G_{\text{diss}}(\text{polar})$, respectively, because their nondominant nonpolar contributions are negligible, and (3) most of $\Delta G_{\text{sol}}(\text{polar})$ and $\Delta G_{\text{diss}}(\text{polar})$ is considered to come from the enthalpic contribution.^{13,45–47}

The last assumption is the entropy–enthalpy compensation,^{48–51} which is expected to effectively hold for the binding of a series of ligands with a protein.

$$T\Delta S = \alpha\Delta H + \text{const (entropy–enthalpy compensation)}$$

where $\alpha > 0$ (usually $1 > \alpha > 0$).

The above three assumptions yield the following expression:

$$\Delta G_{\text{obs}} = A[\Delta E_{\text{bind}} + \Delta G_{\text{sol}}(\text{polar}) + \Delta G_{\text{diss}}(\text{polar})] + \text{const (LERE-QSAR equation)}$$

Coefficient A on the right-hand side of the above equation is a constant determined by α and β [$= (1 - \alpha)(1 + \beta)$].

2.4. Calculation of LERE Representative Energy Terms.

The binding energy $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{EE}) (= E(\text{complex}) - E(\text{protein}) - E(\text{inhibitor}))$ was calculated at the ONIOM(HF/6-31G(d):Amber) level with an electronic embedding (EE) scheme⁵² [the single-point ONIOM(HF/EE) calculation on the optimized structure by the ONIOM(HF/ME) one (HF/EE//HF/ME)]. To account for dispersion forces in the ab initio HF (Hartree–Fock) formalism, the classical Lennard–Jones R^{-6} energy term (LJ6) is added perturbatively to the HF energy ($\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF-D}/\text{EE}) = \Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{EE}) + \text{LJ6}$).^{53–55} The polarization energy $\Delta E_{\text{polarization}}$ was defined as the difference between $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{EE})$ and $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{ME})$.^{56,57} The solvation ($\Delta G_{\text{sol}}(\text{polar})$) and dissociation ($\Delta G_{\text{diss}}(\text{polar})$) energy terms were calculated

Table 3. Binding Interaction Energy^a of MMP-9 with BSs (a), and Overall Free-Energy Change ΔG and Representative Energy Terms^a (b)

(a)									
compound									
no.	−X	ΔE_{bind} (ONIOM/HF-D/EE) ^b	ΔE_{bind} (ONIOM/HF/EE)	LJ6	ΔE_{bind} (MM) ^c	$\Delta E_{\text{bind}}^{\text{ele}}$	$\Delta E_{\text{bind}}^{\text{vdW}}$		
1	OMe	−289.76	−231.56	−58.20	−186.04	−141.41	−44.63		
2	Me	−284.95	−226.69	−58.26	−182.69	−138.87	−43.82		
3	Br	−282.48	−224.15	−58.33	−180.68	−135.84	−44.84		
4	CHO	−280.76	−223.42	−57.34	−178.53	−134.34	−44.19		
5	CF ₃	−282.57	−223.60	−58.96	−181.26	−135.36	−45.90		
6	H	−279.91	−222.58	−57.33	−178.88	−138.13	−40.75		
7	NO ₂	−279.93	−222.20	−57.73	−179.02	−132.34	−46.68		
8	CN	−278.23	−220.65	−57.58	−178.66	−133.30	−45.36		
9	F	−279.26	−221.45	−57.80	−178.19	−137.10	−41.09		
variance ^d		11.3 (3.36) ^e	9.76 (3.12) ^e	0.254 (0.504) ^e	5.94 (2.44) ^e	7.37 (2.71) ^e	3.63 (1.91) ^e		
(b)									
compound									
no.	−X	ΔG_{obs}^f	ΔG_{cal}^g	ΔG_{cal}^h	ΔE_{bind} (ONIOM/HF/EE)	LJ6	$\Delta E_{\text{polarization}}^i$	$\Delta G_{\text{sol}}(\text{polar})$	$\Delta G_{\text{diss}}(\text{polar})^j$
1	OMe	−7.72	−7.16	−7.55	−231.56	−58.20	−6.02	154.94	0.19
2	Me	−7.38	−7.07	−7.43	−226.69	−58.26	−5.35	150.25	0.07
3	Br	−6.95	−7.34		−224.15	−58.33	−4.14	147.18	−0.28
4	CHO	−6.50	−6.27	−6.37	−223.42	−57.34	−4.41	148.58	−0.29
5	CF ₃	−6.41	−6.89		−223.60	−58.96	−4.07	147.57	−0.33
6	H	−6.26	−5.91	−5.88	−222.58	−57.33	−4.85	148.47	0.00
7	NO ₂	−6.03	−6.34	−6.45	−222.20	−57.73	−4.14	147.23	−0.48
8	CN	−5.77	−5.81	−5.74	−220.65	−57.58	−3.57	146.74	−0.36
9	F	−5.71	−5.96	−5.95	−221.45	−57.80	−4.70	147.24	−0.16
variance ^d					9.76 (3.12) ^e	0.254 (0.504) ^e	0.494 (0.703) ^e	5.90 (2.43) ^e	4.44 × 10 ^{−2} (0.211) ^e

^aIn kcal/mol. ^b $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF-D}/\text{EE}) = \Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{EE}) + \text{LJ6}$. ^c $\Delta E_{\text{bind}}(\text{MM}) = \Delta E_{\text{bind}}^{\text{ele}} + \Delta E_{\text{bind}}^{\text{vdW}}$. ^dIn kcal²/mol². ^eA value in parentheses is the standard deviation (in kcal/mol). ^f $\Delta G_{\text{obs}} = 2.303 RT \log IC_{50}$ ($T = 298 \text{ K}$). ^gCalculated from eq 2a. ^hCalculated from eq 2b. ⁱ $\Delta E_{\text{polarization}} = \Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{EE}) - \Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{ME})$. ^jRelative value from $\Delta G_{\text{diss}}(\text{polar})_{\text{X=H}}$. $\Delta G_{\text{diss}}(\text{polar})_{\text{X=H}} = 50.12 \text{ kcal/mol}$.

with PB/SA (Poisson–Boltzmann/surface area)⁵⁸ and SCRFP–IEF-PCM^{59,60} at the HF/6-31+G(d) level, respectively.

3. RESULTS AND DISCUSSION

3.1. Intrinsic Binding Interaction Energy. Table 3a lists the intrinsic binding interaction energy of MMP-9 with each compound, calculated using ONIOM and Amber force field (MM) calculations, $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{EE})$, $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF-D}/\text{EE})$, and $\Delta E_{\text{bind}}(\text{MM})$. Equations 1a and 1b show the nice linearity of $\Delta E_{\text{bind}}(\text{MM})$ with $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{EE})$ and $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF-D}/\text{EE})$. The coefficients of $\Delta E_{\text{bind}}(\text{MM})$ in eqs 1a and 1b are not far from unity.

$$\begin{aligned} \Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{EE}) &= 1.23\Delta E_{\text{bind}}(\text{MM}) - 2.45 \\ n &= 9, r = 0.958, s = 1.02, F = 78.1 \end{aligned} \quad (1a)$$

$$\begin{aligned} \Delta E_{\text{bind}}(\text{ONIOM}/\text{HF-D}/\text{EE}) &= 1.35\Delta E_{\text{bind}}(\text{MM}) - 38.7 \\ n &= 9, r = 0.979, s = 0.768, F = 164.8 \end{aligned} \quad (1b)$$

Although the quality of correlation in eq 1b is slightly better than that in eq 1a, Table 3a shows that the variance of LJ6 among the nine complexes (0.254 kcal²/mol²) is considerably smaller than that of $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{EE})$ (9.76 kcal²/mol²). In spite of differences between quantum and classical treatments, $\Delta E_{\text{bind}}(\text{MM})$ can effectively reproduce $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF}/\text{EE})$ as well as $\Delta E_{\text{bind}}(\text{ONIOM}/\text{HF-D}/\text{EE})$.

Therefore, we decomposed $\Delta E_{\text{bind}}(\text{MM})$ into individual contributions from the i -th amino acid residue ($\Delta E_{\text{bind}}^{i\text{-th residue}}$) in MMP-9 ($\Delta E_{\text{bind}}(\text{MM}) = \sum \Delta E_{\text{bind}}^{i\text{-th residue}}$), and examined the critical interaction that governs the variation of $\Delta E_{\text{bind}}(\text{MM})$. The unit of an amino acid residue in the decomposition is according to the conventional one in biochemistry ($-\text{NH}-\text{C}_\alpha\text{R}-\text{CO}-$).

3.2. Contributions of Individual Amino Acid Residues to the Total Intrinsic Interaction Energy. As shown in Table 3a, the contribution of $\Delta E_{\text{bind}}^{\text{ele}}$ is greater than that of $\Delta E_{\text{bind}}^{\text{vdW}}$ to $\Delta E_{\text{bind}}(\text{MM})$ ($= \Delta E_{\text{bind}}^{\text{ele}} + \Delta E_{\text{bind}}^{\text{vdW}}$), because the variance of $\Delta E_{\text{bind}}^{\text{ele}}$ (7.37 kcal²/mol²) is approximately twice larger than that of $\Delta E_{\text{bind}}^{\text{vdW}}$ (3.63 kcal²/mol²). This suggests that the complex formation is driven predominantly by the electrostatic interaction. Figure 3a shows the contribution of each amino acid residue in the complex between compound 6 ($\text{X} = \text{H}$) and the i -th residue in MMP-9 to the overall intrinsic binding energy $\Delta E_{\text{bind}}(\text{MM})$. Large absolute values of $\Delta E_{\text{bind}}^{i\text{-th residue}}$ found in Figure 3a come from the strong electrostatic interaction of charged amino acid residues with the negatively charged compound 6, and negatively charged residues (Asp and Glu) exhibit repulsive interaction energy, whereas positively charged ones (Arg and Lys), two zinc ions, and four calcium ions exhibit attractive interaction energy. These charged residues and ions have a large influence on $\Delta E_{\text{bind}}(\text{MM})$, because of the long-distance nature of electrostatic interaction. As expected, the zinc ion in the active site plays a decisive role in the binding (−166 kcal/mol). Glu402 and the water molecule in the active site (see Figure 2) undergo hydrogen-bonding and/or electrostatic interactions with the ligand (−14.6 and −8.09 kcal/mol,

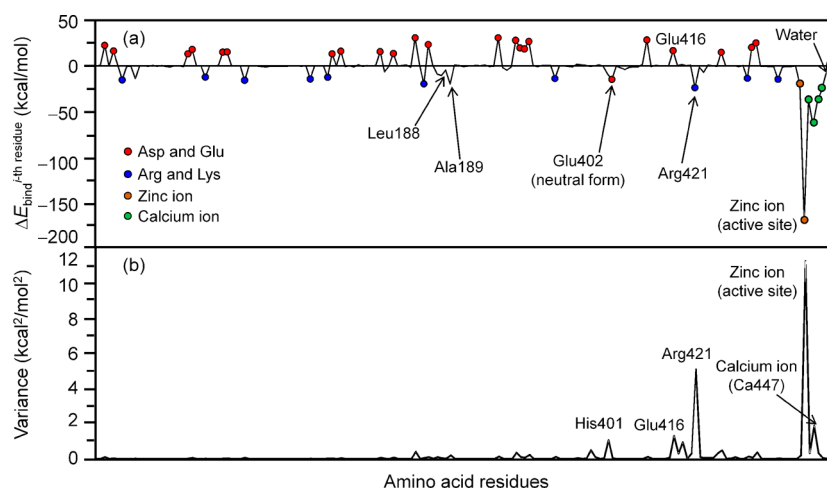


Figure 3. Interaction energy profile of $\Delta E_{\text{bind}}(\text{MM})$. (a) $\Delta E_{\text{bind}}^{i\text{-th residue}}(\Delta E_{\text{bind}}(\text{MM}))$ of compound 6 ($X = \text{H}$) with the i -th amino acid residue in MMP-9. (b) Variance of $\Delta E_{\text{bind}}^{i\text{-th residue}}$ among nine BS–MMP-9 complexes.

respectively). Similarly, Leu188 and Ala189 undergo relatively weak interactions (-10.2 and -4.4 kcal/mol, respectively).

Figure 3b shows the variance of $\Delta E_{\text{bind}}^{i\text{-th residue}}$ among the total nine complexes. The zinc ion in the active site exhibits the greatest variance (10.0 kcal 2 /mol 2), clearly revealing that the variation in the overall intrinsic binding energy is determined mostly by the interaction energy of the zinc ion as to the carboxylate group in compounds 1–9. Arg421 shows the second largest variance (4.54 kcal 2 /mol 2). This is probably because the charged side chain atom of Arg421 is located near the substituent (X) in compounds 1–9 (nearest to X among the six Arg residues in MMP-9), although the smallest distance between them is larger than the sum of the van der Waals radii of the nearest neighboring atoms (>4 Å). As a result, a subtle change in the distance causes a large change in the interaction energy, leading to the large variance of Arg421. For a similar reason, Glu416 and the calcium ion (Ca447) closest to the active site among the four calcium ions show relatively large variances (1.16 and 1.69 kcal 2 /mol 2 , respectively). It should be noted that the variances of Leu188, Ala189, Glu402, and the water molecule, which undergo electrostatic/hydrogen-bonding interactions directly with a ligand, are 2.80×10^{-2} , 1.81×10^{-2} , 2.65×10^{-2} , and 7.30×10^{-3} kcal 2 /mol 2 , respectively, showing that these residues make only minor contributions to the variance of $\Delta E_{\text{bind}}(\text{MM})$.

3.3. LERE-QSAR Analysis. As explained in section 3.1, $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ is nicely reproducible with $\Delta E_{\text{bind}}(\text{MM})$, because of the good linearity between these two terms. Here, we propose a QSAR model involving the LERE representative energy terms including quantum-chemically derived $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$, which is supposed to be more adequate and reliable than the classically derived one, $\Delta E_{\text{bind}}(\text{MM})$, for describing the intermolecular interaction.

Table 3b shows the representative energy terms, $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$, LJ6, $\Delta E_{\text{polarization}}$, $\Delta G_{\text{sol}}(\text{polar})$, and $\Delta G_{\text{diss}}(\text{polar})$, together with their variance. The variances of $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ and $\Delta G_{\text{sol}}(\text{polar})$ are considerably large (9.76 and 5.90 kcal 2 /mol 2 , respectively). As mentioned previously, the variance of LJ6 is much smaller (0.256 kcal 2 /mol 2). Those of $\Delta E_{\text{polarization}}$ and $\Delta G_{\text{diss}}(\text{polar})$ are also smaller (0.494 and 4.44×10^{-2} kcal 2 /mol 2 , respectively). Therefore, variations of $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ and $\Delta G_{\text{sol}}(\text{polar})$ are expected to explain that of ΔG_{obs} quantitatively.

We performed the LERE-QSAR analysis using the two terms $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ and $\Delta G_{\text{sol}}(\text{polar})$. Although eq 2a for compounds 1–9 is statistically significant, the quality is not so good.

$$\Delta G_{\text{obs}} = 0.500[\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE}) + \Delta G_{\text{sol}}(\text{polar})] + 31.2$$

$$n = 9, r = 0.848, s = 0.396, F = 17.9 \quad (2a)$$

We found that there is a nice anticorrelation between ΔE_{bind} and $\Delta G_{\text{sol}}(\text{polar})$ for the binding of a congeneric series of ligands with their target protein.^{10,11,13,44} In the current case, an acceptable anticorrelation between $\Delta G_{\text{sol}}(\text{polar})$ and $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ was observed ($n = 9, r = -0.950$), as can be seen in Figure 4.

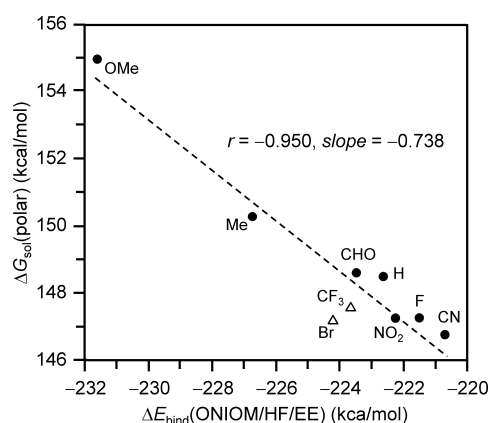


Figure 4. Plots of $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ with $\Delta G_{\text{sol}}(\text{polar})$.

This anticorrelation means that stabilization of the overall free-energy by $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ counteracts destabilization by $\Delta G_{\text{sol}}(\text{polar})$. Compounds 3 ($X = \text{Br}$) and 5 ($X = \text{CF}_3$) show relatively large deviation from the regression line in Figure 4. Although there are several conceivable reasons for these “outliers”, one of the reasons may be a lack of accuracy of $\Delta G_{\text{sol}}(\text{polar})$ and $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ for estimating the very subtle energy change (the variation ranges of $\Delta G_{\text{sol}}(\text{polar})$ and $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ are 150.84 ± 4.10 and -226.11 ± 5.45 kcal/mol, respectively). Removal of compounds 3 and 5 improves the statistical quality ($n = 7, r = -0.988$).

Accordingly, removal of compounds 3 and 5 from eq 2a yields eq 2b.

$$\begin{aligned}\Delta G_{\text{obs}} &= 0.668[\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE}) \\ &\quad + \Delta G_{\text{sol}}(\text{polar})] + 43.6 \\ n &= 7, r = 0.940, s = 0.293, F = 38.1, \\ \text{outliers} &= \text{compounds 3 and 5}\end{aligned}\quad (2b)$$

Comparison between variations of $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ and $\Delta G_{\text{sol}}(\text{polar})$ indicates that the contribution from the former overwhelms that from the latter, as can be seen in Figure 5. Our

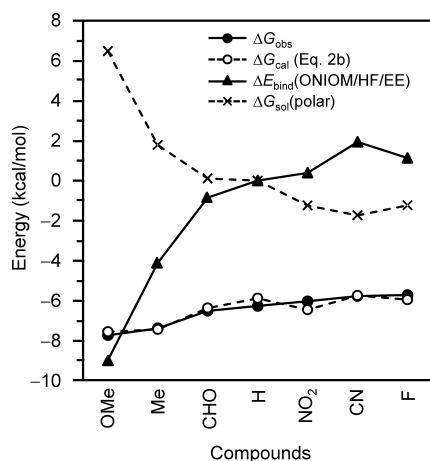


Figure 5. Change in free-energy components. $\Delta E_{\text{bind}}(\text{OMIOM}/\text{HF}/\text{EE})$ and $\Delta G_{\text{sol}}(\text{polar})$ are relative energies from those of compound 6 ($X = \text{H}$).

previous results obtained through the LERE-QSAR analysis of a series of benzenesulfonamides with CA (carbonic anhydrase), which is also a zinc-containing protease, suggest that the contribution from $\Delta G_{\text{sol}}(\text{polar})$ and $\Delta G_{\text{diss}}(\text{polar})$ overwhelms that from $\Delta E_{\text{bind}}(\text{FMO}/\text{HF}/6\text{-}31\text{G})$.^{10,11} It is noteworthy that ΔE_{bind} governs ΔG_{obs} in the present case, unlike the case of CA.

To summarize this section, the binding interaction energy of the carboxylate group (in a ligand) with the zinc ion in the active site of MMP-9 decisively governs the variation of the inhibitory potency.

4. CONCLUSION

As mentioned at the beginning, both MMP-9 and CA (carbonic anhydrase) belong to the zinc-containing protease family. However, classical QSAR equations for two series of ligand–protein complexes, MMP-9–BSs and CA–benzenesulfonamides, appear to exhibit different dependence on the Hammett σ constant of the substituent in ligands. While the pIC_{50} of BSs with MMP-9 is negatively correlated with the σ ($\text{pIC}_{50} [= -\Delta G_{\text{obs}}/(2.303 RT)] = a\sigma + bMR + \text{const}$, $a < 0$), the $\text{pK}_i [= \log(1/K_i) = -\Delta G_{\text{obs}}/(2.303 RT)]$ of a series of benzenesulfonamides with CA is positively correlated ($\text{pK}_i = a\sigma + b\log P + \text{const}$, $a > 0$).⁶¹ Comparison of the present and previous LERE-QSAR results^{10,11} for the two sets reveals differences in the relative contributions of the LERE representative energy terms to ΔG_{obs} ; the contributions from $\Delta G_{\text{sol}}(\text{polar})$ and $\Delta G_{\text{diss}}(\text{polar})$ in the two cases of CA and MMP-9 are larger and smaller than that of $\Delta E_{\text{bind}}(\text{MO})$, obtained with ab initio MO calculations (FMO and ONIOM), respectively. It should be noted that the former two and latter representative energy terms in both cases exhibit

nice negative and positive correlations with ΔG_{obs} , respectively. Probably, the degree of sensitivity of structural changes in the two congeneric series of ligands to each of electrostatic representative energy terms, $\Delta G_{\text{sol}}(\text{polar})$, $\Delta G_{\text{diss}}(\text{polar})$, and $\Delta E_{\text{bind}}(\text{MO})$, determines their relative contributions, independently.

The interaction of a ligand with the zinc ion both in MMP-9 and CA contributes extremely significantly to the variation of the intrinsic binding energy (ΔE_{bind}) among ligands, as well as the stabilization of the total binding free-energy of each ligand. However, it should be noted that the structural factors that stabilize the total binding free-energy are sometimes different from those that govern the variation in the total binding free-energy among a series of ligands.¹²

Although results such as those discussed in the present work are probably unobtainable solely with a classical QSAR type of approach, the current LERE-QSAR analysis has the ability to provide physicochemically interpretable and quantitative explanations.

Although molecular orbital studies concerning an enzymatic reaction have provided basic and useful information on the electronic and atomic mechanisms,^{27–37,56,57,62,63} a combination of molecular orbital and QSAR approaches,^{10–13,64–66} such as one shown in this work, can provide different and additional information. For understanding variation in binding affinity among ligands, it is advantageous to carefully examine electronic and steric changes in the interaction of a protein with a congeneric series of ligands, using molecular calculations on whole complex structures, because a congeneric series of ligands are expected to often exhibit a “certain systematic” effect on the interaction energy with their target protein, as suggested by a number of QSAR studies so far.

■ ASSOCIATED CONTENT

Supporting Information

Tables S1-1–S1-9 that list interatomic distances of all of the compounds 1–9. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: +81-88-633-7257. Fax: +81-88-633-9508. E-mail: hchuman@tokushima-u.ac.jp.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by Grants-in-Aid for Scientific Research (Nos. 24590051 and 23710272) from the Ministry of Education, Culture, Sports, Science and Technology. We thank Emeritus Professor Toshio Fujita (Kyoto University, Japan) for his instructive suggestions.

■ REFERENCES

- (1) Andreini, C.; Banci, L.; Bertini, I.; Rosato, A. *J. Proteome Res.* **2006**, *5*, 196–201.
- (2) Lipscomb, W. N.; Sträter, N. *Chem. Rev.* **1996**, *96*, 2375–2433.
- (3) Gupta, S. P. *Chem. Rev.* **2007**, *107*, 3042–3087.
- (4) Maret, W.; Li, Y. *Chem. Rev.* **2009**, *109*, 4682–4707.
- (5) Corradi, H. R.; Schwager, S. L. U.; Nchinda, A. T.; Sturrock, E. D.; Acharya, K. R. *J. Mol. Biol.* **2006**, *357*, 964–974.
- (6) Christianson, D. W.; Lipscomb, W. N. *Acc. Chem. Res.* **1989**, *22*, 62–69.

- (7) Matthews, B. W. *Acc. Chem. Res.* **1988**, *21*, 333–340.
- (8) Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. *Chem. Rev.* **2008**, *108*, 946–1051.
- (9) Anzellotti, A. I.; Farrell, N. P. *Chem. Soc. Rev.* **2008**, *37*, 1629–1651.
- (10) Yoshida, T.; Munei, Y.; Hitaoka, S.; Chuman, H. *J. Chem. Inf. Model.* **2010**, *50*, 850–860.
- (11) Munei, Y.; Shimamoto, K.; Harada, M.; Yoshida, T.; Chuman, H. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 141–144.
- (12) Hitaoka, S.; Harada, M.; Yoshida, T.; Chuman, H. *J. Chem. Inf. Model.* **2010**, *50*, 1796–1805.
- (13) Hitaoka, S.; Matoba, H.; Harada, M.; Yoshida, T.; Tsuji, D.; Hirokawa, T.; Itoh, K.; Chuman, H. *J. Chem. Inf. Model.* **2011**, *51*, 2706–2716.
- (14) Overall, C. M. *Mol. Biotechnol.* **2002**, *22*, 51–86.
- (15) Parks, W. C.; Wilson, C. L.; López-Boado, Y. S. *Nat. Rev. Immunol.* **2004**, *4*, 617–629.
- (16) Galis, Z. S.; Khatri, J. J. *Circ. Res.* **2002**, *90*, 251–262.
- (17) Creemers, E. E.; Cleutjens, J. P.; Smits, J. F.; Daemen, M. J. *Circ. Res.* **2001**, *89*, 201–210.
- (18) Verma, R. P.; Hansch, C. *Bioorg. Med. Chem.* **2007**, *15*, 2223–2268.
- (19) O'Brien, P. M.; Ortwine, D. F.; Pavlovsky, A. G.; Picard, J. A.; Sliskovic, D. R.; Roth, B. D.; Dyer, R. D.; Johnson, L. L.; Man, C. F.; Hallak, H. J. *Med. Chem.* **2000**, *43*, 156–166.
- (20) Tochowicz, A.; Maskos, K.; Huber, R.; Oltenfreiter, R.; Dive, V.; Yiotakis, A.; Zanda, M.; Bode, W.; Goettig, P. *J. Mol. Biol.* **2007**, *371*, 989–1006.
- (21) Johnson, L. L.; Pavlovsky, A. G.; Johnson, A. R.; Janowicz, J. A.; Mani, C. F.; Ortwine, D. F.; Purchase, C. F.; White, A. D.; Hupe, D. J. *J. Biol. Chem.* **2000**, *275*, 11026–11033.
- (22) Dolinsky, T. D.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. *Nucleic Acids Res.* **2004**, *32*, W665–W667.
- (23) Sbardella, D.; Fasciglione, G. F.; Gioia, M.; Ciaccio, C.; Tundo, G. R.; Marini, S.; Coletta, M. *Mol. Aspects Med.* **2012**, *33*, 119–208.
- (24) Frisch, M. J.; et al. *Gaussian 09*, revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.
- (25) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (26) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357–19363.
- (27) Lundberg, M.; Morokuma, K. *J. Phys. Chem. B* **2007**, *111*, 9380–9389.
- (28) Lundberg, M.; Kawatsu, T.; Vreven, T.; Frisch, M. J.; Morokuma, K. *J. Chem. Theory Comput.* **2009**, *5*, 222–234.
- (29) Li, X.; Chung, L. W.; Paneth, P.; Morokuma, K. *J. Am. Chem. Soc.* **2009**, *131*, 5115–5125.
- (30) Ananikov, V. P.; Musaev, D. J.; Morokuma, K. *J. Mol. Catal. A: Chem.* **2010**, *324*, 104–119.
- (31) Lundberg, M.; Sasakura, Y.; Zheng, G.; Morokuma, K. *J. Chem. Theory Comput.* **2010**, *6*, 1413–1427.
- (32) Sekharan, S.; Yokoyama, S.; Morokuma, K. *J. Phys. Chem. B* **2011**, *115*, 15380–15388.
- (33) Ke, Z.; Abe, S.; Ueno, T.; Morokuma, K. *J. Am. Chem. Soc.* **2011**, *133*, 7926–7941.
- (34) Sekharan, S.; Morokuma, K. *J. Am. Chem. Soc.* **2011**, *133*, 19052–19055.
- (35) Barone, V.; De Rienzo, F.; Langella, E.; Menziani, M. C.; Rega, N.; Sola, M. *Proteins* **2006**, *62*, 262–269.
- (36) Pelmenchikov, V.; Siegbahn, P. E. M. *Inorg. Chem.* **2002**, *41*, 5659–5666.
- (37) Tao, P.; Fisher, J. F.; Shi, Q.; Mobashery, S.; Schlegel, H. B. *J. Phys. Chem. B* **2010**, *114*, 1030–1037.
- (38) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
- (39) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
- (40) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- (41) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (42) Bren, U.; Martínek, V.; Florián, J. *J. Phys. Chem. B* **2006**, *110*, 12782–12788.
- (43) Bren, M.; Florián, J.; Mavri, J.; Bren, U. *Theor. Chem. Acc.* **2007**, *117*, 535–540.
- (44) Yoshida, T.; Shimizu, M.; Harada, M.; Hitaoka, S.; Chuman, H. *Bioorg. Med. Chem. Lett.* **2011**, *22*, 124–128.
- (45) Noyes, R. M. *J. Am. Chem. Soc.* **1962**, *84*, 513–522.
- (46) Rashin, A. A.; Honig, B. *J. Phys. Chem.* **1985**, *89*, 5588–5593.
- (47) Chuman, H.; Mori, A.; Tanaka, H.; Yamagami, C.; Fujita, T. *J. Pharm. Sci.* **2004**, *93*, 2681–2697.
- (48) Dunitz, J. D. *Chem. Biol.* **1995**, *2*, 709–712.
- (49) Gallicchio, E.; Kubo, M. M.; Levy, R. M. *J. Am. Chem. Soc.* **1998**, *120*, 4526–4527.
- (50) Sharp, K. *Protein Sci.* **2001**, *10*, 661–667.
- (51) Freire, E. *Chem. Biol. Drug Des.* **2009**, *74*, 468–472.
- (52) Vreven, T.; Byun, K. S.; Komromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.
- (53) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (54) Gonzalez, C.; Lim, E. C. *J. Phys. Chem. A* **2003**, *107*, 10105–10110.
- (55) He, X.; Fusti-Molnar, L.; Cui, G.; Merz, K. M., Jr. *J. Phys. Chem. B* **2009**, *113*, 5290–5300.
- (56) Hirao, H. *J. Phys. Chem. B* **2011**, *115*, 11278–11285.
- (57) Hirao, H. *Chem. Lett.* **2011**, *40*, 1179–1181.
- (58) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; et al. *Acc. Chem. Res.* **2000**, *33*, 889–897.
- (59) Tomasi, J. *Theor. Chem. Acc.* **2004**, *112*, 184–203.
- (60) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (61) Hansch, C.; McClarin, J.; Klein, T.; Langridge, R. *Mol. Pharmacol.* **1985**, *27*, 493–498.
- (62) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. *Chem. Rev.* **2012**, *112*, 632–672.
- (63) Mazanetz, M. P.; Ichihara, O.; Law, R. J.; Whittaker, M. *J. Cheminf.* **2011**, *3*, 2.
- (64) Yoshida, T.; Yamagishi, K.; Chuman, H. *QSAR Comb. Sci.* **2008**, *27*, 694–703.
- (65) Yoshida, T.; Fujita, T.; Chuman, H. *Curr. Comput.-Aided Drug Des.* **2009**, *5*, 38–55.
- (66) Yoshida, T.; Hirozumi, K.; Harada, M.; Hitaoka, S.; Chuman, H. *J. Org. Chem.* **2011**, *76*, 4564–4570.