

# Binding Free-Energy Calculation Is a Powerful Tool for Drug Optimization: Calculation and Measurement of Binding Free Energy for 7-Azaindole Derivatives to Glycogen Synthase Kinase-3 $\beta$

Kunihiro Kitamura,<sup>\*,†</sup> Yunoshin Tamura,<sup>‡</sup> Tomokazu Ueki,<sup>‡</sup> Koji Ogata,<sup>§</sup> Shigeho Noda,<sup>§</sup> Ryutaro Himeno,<sup>§</sup> and Hiroshi Chuman<sup>||</sup>

<sup>†</sup>Nihon Pharmaceutical University, 10281 Komuro, Ina-cho, Kitaadachi-gun, Saitama 362-0806, Japan

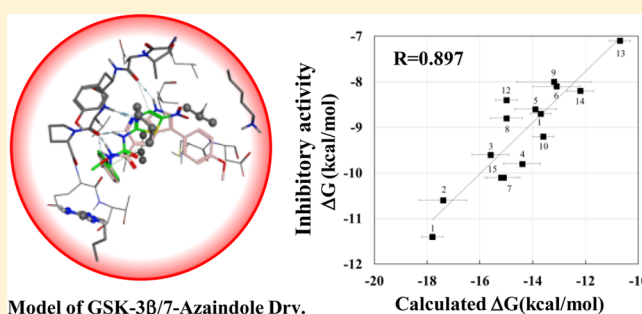
<sup>‡</sup>Taisho Pharmaceutical Co. Ltd., 1-403 Yoshino-cho, Kita-ku, Saitama-shi, Saitama 331-9530, Japan

<sup>§</sup>RIKEN, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan

<sup>||</sup>Institute of Health Biosciences, University of Tokushima Graduate School, 1-78 Shomachi, Tokushima, Tokushima 770-8505, Japan

## Supporting Information

**ABSTRACT:** Present computational lead (drug)-optimization is lacking in thermodynamic tactics. To examine whether calculation of binding free-energy change ( $\Delta G$ ) is effective for the lead-optimization process, binding  $\Delta G$ s of 7-azaindole derivatives to the ATP binding site of glycogen synthase kinase-3 $\beta$  (GSK-3 $\beta$ ) were calculated. The result was a significant correlation coefficient of  $r = 0.895$  between calculated and observed  $\Delta G$ s. This indicates that calculated  $\Delta G$  reflects the inhibitory activities of 7-azaindole derivatives. In addition to quantitative estimation of activity,  $\Delta G$  calculation characterizes the thermodynamic behavior of 7-azaindole derivatives, providing also useful information for inhibitor optimization on affinity to water molecules.



## ■ INTRODUCTION

Development of new effective medicines starts from discovery of lead compounds that bind to target protein active sites through the biological screening system. The important thing at this screening stage is to obtain a significant number of lead compounds for different chemical structures because a small quantity of lead compounds does not always lead to a drug molecule with sufficient biological activity and solubility. The high-throughput screening (HTS) technique enables rapid discovery of many lead candidates from a huge number of compounds. For this reason, HTS has been widespread among all pharmaceutical companies at present.

However, no effective approach has yet been developed successfully for the subsequent lead-optimization stage. Indeed, a major technique that has been applied for this stage, molecular calculations<sup>1,2</sup> based on 3-dimensional (3D) models of drug candidates docking to the target protein, has provided a variety of structural information such as hydrogen bonding formation, electrostatic interaction, molecular surface complementarity, and so on. In most cases, however, candidates designed by using such information have hardly ever shown high activity together with enough solubility to be developed as a medical drug. We believe this is largely due to the following reason. The 3D docking model has not been well examined from the standpoint of

thermodynamics even though the actual binding progresses under free-energy change ( $\Delta G$ ) as shown in formula 1



where P, L, and PL denote target protein, drug molecule, and drug–protein complex, respectively, and subscript “aq” means that P, L, and PL exist as an aqueous solution. Efficient success of drug optimization therefore depends on whether  $\Delta G$  can be estimated as exactly and easily as possible. Statistical mechanics approaches<sup>3–5</sup> to obtaining binding  $\Delta G$  have already been performed, and several calculation methods are widely used such as thermodynamics integration (TI),<sup>3</sup> free energy perturbation (FEP),<sup>4</sup> and Bennett’s acceptance ratio (BAR).<sup>5</sup> Direct calculation of  $\Delta G$  itself in formula 1 is not able to guarantee exact estimation because of restrictions on available computational resources. To calculate binding  $\Delta G$  as exactly as possible, we thus adopted a unique approach that uses the double annihilation (DA)<sup>6,7</sup> technique in accordance with the following scheme

$$\Delta G = \Delta G_{\text{complex}} - \Delta G_{\text{sol}} \quad (2)$$



Received: December 6, 2013

$$L_{\text{gas}} \xrightarrow{\Delta G_{\text{sol}}} L_{\text{aq}} \quad (4)$$

where  $\Delta G_{\text{complex}}$  and  $\Delta G_{\text{sol}}$  respectively represent free-energy change of the complex-formation in formula 3 and the free-energy change of solvation in formula 4. Here, it is noteworthy that formula 2 reflects the thermodynamic behavior of an inhibitor molecule, as mentioned below.

In this paper, we report how effectively this  $\Delta G$  calculation approach works for quantitative estimation of the inhibitory activity of drug candidates and also for the drug design of promising inhibitors through the example of GSK-3 $\beta$  and its inhibitors, 7-azaindole derivatives.

GSK-3 $\beta$ , a multifunctional serine/threonine kinase, is a key regulator of numerous signaling pathways.<sup>8</sup> It inactivates glycogen synthase<sup>9</sup> and also phosphorylates<sup>10,11</sup> tau protein relevant to Alzheimer's disease (AD). It is thus in the spotlight<sup>12–16</sup> as a target protein for new drug developments for type-2 diabetes and AD.

## MATERIALS AND METHODS

**Generation of 3D Docking Structure Model.** Inhibitor molecules for protein kinases are expected as drug candidates for a number of diseases including cancer, diabetes, and dementia, as mentioned above. Thus, crystal structures of many protein kinases with their inhibitors have been elucidated (see Protein Data Bank<sup>17</sup>). Any crystal structures of GSK-3 $\beta$  with the derivatives in Table 1, however, have not been reported. It has been reported that 7-azaindole binds<sup>18</sup> to the ATP binding site in a chimera protein of protein kinase A and B (PKA–PKB chimera), interacting through hydrogen bonds of heterocyclic N and NH to the backbone amides of the chimera protein. In addition, a crystal structure<sup>15</sup> of GSK-3 $\beta$  with inhibitor AR-A014418 (Figure 1) showed that the inhibitor interacts through three hydrogen bonds with backbone amides of Val135 and Pro136 (PDB-based number) in the ATP binding site. Such a pattern of hydrogen bonding to Val135 is similar to that between 7-azaindole and PKA–PKB chimera.

Accordingly, we assumed that 7-azaindole derivatives bind to the ATP-binding site of GSK-3 $\beta$  and superimposed 15 derivatives on AR-A014418 in the crystal structure of the complex, as hydrogen bonds are formed between heterocyclic N atoms and the backbone atoms. Then, an energy minimum structure of each modeled complex was selected as the initial structure for equilibration calculation with molecular dynamics (MD) simulation.

**Generation of the Molecular System for Equilibration Calculation.** To perform free-energy calculations by using DA, each derivative and its complex with GSK-3 $\beta$  were constructed in an aqueous phase by using the MOE molecular modeling package (Ver. 2008.10).<sup>19</sup> Water molecules were placed to form a truncated octahedron box with a periodic boundary condition. TIP3P water molecules<sup>20</sup> were added. Around 1,000 water molecules were added to the compounds and around 12,000 water molecules to the GSK-3 $\beta$ /compound complex. Since a neutral condition was required for Particle-Mesh Ewald (PME)<sup>21</sup> calculations, we added 9 chloride ions to each GSK-3 $\beta$ /compound complex. Force-field parameters of ff99SB<sup>22</sup> and GAFF<sup>23</sup> with RESP (HF/6-31G\*) charges<sup>24</sup> were used. These parameters were assigned by using AMBER 9 modules<sup>25</sup> and Gaussian 03 software.<sup>26</sup>

To equilibrate each derivative and its complex, energy minimization and MD calculations were performed by using the Gromacs 4.5.3 software suite.<sup>27</sup> First, 10,000 steepest-descent

energy minimization steps were carried out. In the first half of the minimization steps, position restraints were applied to heavy atoms in each complex and compound in order to relax the positions of surrounding water molecules. Second, after a 0.2 ns MD simulation with the same restraints, we performed 5 and 20 ns equilibration runs for each compound and complex, respectively. During the simulation, the temperature was maintained at 298 K by a stochastic dynamics algorithm, and pressure was maintained at 1 bar by scaling the box size in accordance with the Berendsen algorithm.<sup>28</sup> The frictional coefficient and time step were taken as  $\gamma = 5 \text{ ps}^{-1}$  and 2.0 fs, respectively. Snapshot structures were extracted from equilibration runs as initial structures for successive annihilation MD simulations.

**Free Energy Calculation.** The binding free-energy was calculated by using DA and BAR.<sup>29–31</sup> We examined the reverse processes expressed by eqs 3 and 4. Each compound was transferred from aqueous to gas phases. A “partially annihilated state” (window) was introduced to compute the two processes expressed by eqs 3 and 4. The annihilation of free energy was obtained as the sum of free-energy differences between the neighboring windows. In each window, the intermolecular interaction of a compound with GSK-3 $\beta$  and water molecules was described by using soft-core potentials including the annihilation parameters  $\lambda_C$  and  $\lambda_{LJ}$  for Coulomb and Lennard Jones interactions, respectively. The free-energy difference between the  $i$ -th and  $(i+1)$ -th windows was calculated by using BAR. “Non-equilibrium work” ( $W = U_{i+1} - U_i$ , where  $U_i$  represents the potential energy of the  $i$ -th window) was sampled in the MD trajectory. According to the maximum likelihood argument of BAR,<sup>29</sup> the free-energy estimator for two simulations of equal length can be written as

$$\begin{aligned} & \langle [1 + \exp(\beta W - \beta \Delta g_i)]^{-1} \rangle_i \\ &= \langle [1 + \exp(-\beta W + \beta \Delta g_i)]^{-1} \rangle_{i+1} \end{aligned} \quad (5)$$

where  $\Delta g_i$  is the annihilation free-energy change from the  $i$ -th window to the  $(i+1)$ -th window ( $\Delta G = \Sigma \Delta g_i$ ), the subscripted angle bracket indicates an ensemble average over each window, and  $\beta$  is the reciprocal temperature.

Annihilating simulations were independently performed at  $\lambda_C = (0.000, 0.100, 0.250, 0.400, 0.550, 0.650, 0.725, 0.800, 0.875, 0.950, 1.000)$  with  $\lambda_{LJ} = 0$  and at  $\lambda_{LJ} = (0.000, 0.100, 0.200, 0.300, 0.400, 0.475, 0.550, 0.600, 0.650, 0.675, 0.700, 0.725, 0.750, 0.775, 0.800, 0.825, 0.850, 0.875, 0.900, 0.950, 1.000)$  with  $\lambda_C = 1$  (32  $\lambda$  points).<sup>31</sup> No restraint energy was applied to bind a compound to the binding pocket. A 10,000-step steepest descent minimization and a 0.2 ns equilibration of water molecules were carried out for each window. After a 0.2 ns MD simulation for equilibration, 2.5 and 1.0 ns runs for data collection were performed for each hydrated complex and compound, respectively. “Non-equilibrium work” between neighboring windows was measured every 0.1 ps. Other conditions for each simulation were the same for the previous long equilibrating run. To reduce artifacts due to the initial structure selection and to improve the statistical reliability, we performed 4 or 12 independent sets of annihilation simulations.

**Additional 10 ns MD Simulation to Examine Docking Structure of 7-Azaindole Derivative.** Additional 10 ns MD simulations on 15 GSK-3 $\beta$ /derivative complexes were performed to (i) examine relationships between inhibitory activities and docking structures and (ii) compare the results in this study with those of other methods; MMPB/SA, *Ligscore*,<sup>32</sup> and PMF04.<sup>33</sup>

Table 1. Chemical Structures of 7-Azaindole Derivatives and Their Properties<sup>a</sup>

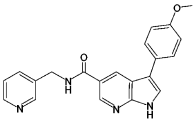
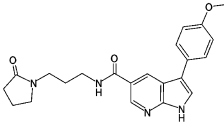
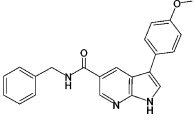
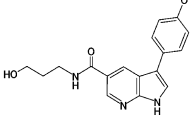
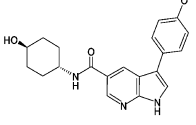
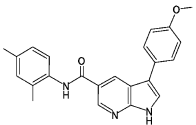
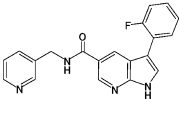
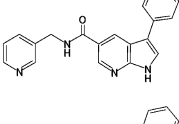
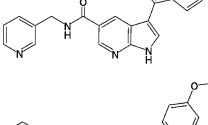
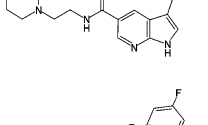
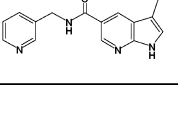
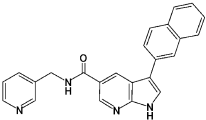
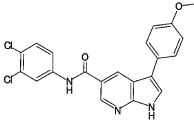
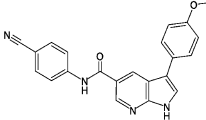
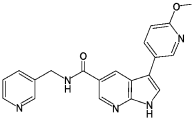
Derivative Number	Chemical Structure	IC <sub>50</sub> <sup>*1</sup> (nM)	$\Delta G_{\text{obs}}$ <sup>*2</sup> (kcal/mol)	$\Delta G$ <sup>*3</sup> (kcal/mol)	$\Delta G_{\text{complex}}$ <sup>*3</sup> (kcal/mol)	$\Delta G_{\text{sol}}$ <sup>*3</sup> (kcal/mol)
1		5	-11.4	-17.8±0.4	-38.8±0.4	-21.1±0.1
2		17	-10.6	-17.4±0.9	-40.5±0.8	-23.1±0.5
3		96	-9.6	-15.6±0.7	-33.7±0.7	-18.1±0.1
4		63	-9.8	-14.4±0.7	-36.8±0.7	-22.4±0.3
5		500	-8.6	-13.9±0.8	-37.1±0.8	-23.1±0.1
6		1206	-8.1	-13.1±0.9	-30.7±0.9	-17.6±0.1
7		43	-10.1	-15.1±0.6	-35.2±0.6	-20.2±0.1
8		326	-8.8	-15.0±0.6	-34.8±0.6	-19.8±0.1
9		1387	-8.0	-13.2±1.4	-37.8±1.4	-24.7±0.1
10		192	-9.2	-13.6±0.4	-33.9±0.3	-20.4±0.1
11		445	-8.7	-13.7±0.4	-33.4±0.3	-19.7±0.1

Table 1. continued

Derivative Number	Chemical Structure	IC <sub>50</sub> <sup>*1</sup> (nM)	ΔG <sub>obs</sub> <sup>*2</sup> (kcal/mol)	ΔG <sup>*3</sup> (kcal/mol)	ΔG <sub>complex</sub> <sup>*3</sup> (kcal/mol)	ΔG <sub>sol</sub> <sup>*3</sup> (kcal/mol)
12		693	-8.4	-15.0±0.4	-36.1±0.4	-21.1±0.1
13		5745	-7.1	-10.7±0.4	-28.6±0.3	-17.9±0.1
14		1054	-8.2	-12.2±0.5	-32.8±0.5	-20.6±0.1
15		42	-10.1	-15.2±0.6	-37.4±0.6	-22.3±0.1

<sup>a</sup>\*1: IC<sub>50</sub> means 50% inhibitory concentration and was measured twice. The averaged value is shown in the table. \*2: ΔG<sub>obs</sub> means observed ΔG calculated by ΔG<sub>obs</sub> = RT ln (IC<sub>50</sub>), where R is the gas constant and T is the absolute temperature. \*3: ΔG, ΔG<sub>complex</sub>, and ΔG<sub>sol</sub> are defined by formulas (2), (3), and (4), respectively; 12 trials of ΔG calculations were performed on 8 derivatives from 1 to 8, and 4 trials for the remaining 7 derivatives from 9 to 15.

The GSK-3β/derivative complexes were optimized by using the conjugate-gradient method (part of Amber 9) in 10,000 steps for hydrogen atoms after 10,000 steps for all atoms. After the optimization, a 50 ps MD simulation under constant volume (NVT) was performed with constraint for the GSK-3β/derivative complex molecule. To obtain an equilibrium system, a 1 ns MD simulation was performed under constant pressure (1 atm) and temperature (T = 300 K) (NPT) by using a Langevin thermostat with the collision frequency γ = 1. In this procedure, anisotropic Berendsen control was used with a time constant of 1 ps. From the initial structure, MD simulation was started at T = 300 K and performed for 10 ns. The snapshots were stored every 10 ps, with 1000 snapshots in total.

**MMPB/SA and MMGB/SA Calculations.** It excited our curiosity to know ΔG values were calculated by other methods for 7-azaindole derivatives. Thus, we performed molecular mechanics Poisson–Boltzmann surface area (MMPB/SA) and molecular mechanics Generalized Born surface area (MMGB/SA) methods<sup>34</sup> in accordance with the following formula 6

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\Delta S \quad (6)$$

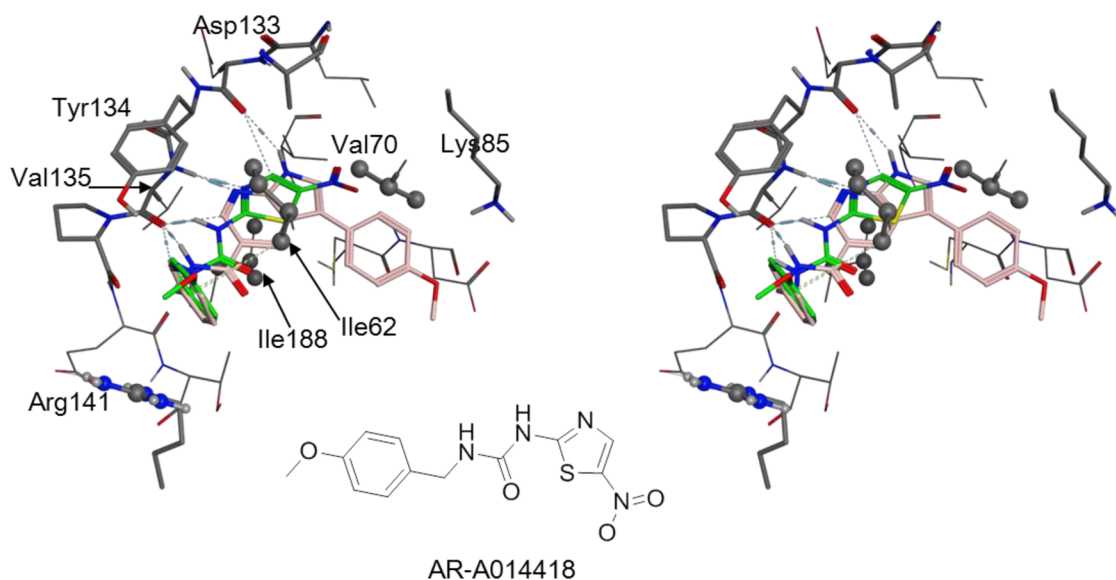
where ΔE<sub>MM</sub>, ΔG<sub>solv</sub>, and -TΔS are the molecular mechanics (MM) energy, the solvation free energy, and the conformational entropy, respectively. To calculate ΔE<sub>MM</sub> and ΔG<sub>solv</sub>, 100 snapshots were extracted from stored structures every 100 ps during 10 ns MD simulation mentioned above. Normal-mode analysis was applied to calculate the conformational entropy.

**GSK-3β Inhibition.** Standard assay for GSK-3β enzyme activity was carried out in duplicates in the presence of different concentrations of test compounds in 30 mM MOPS/NaOH, pH 7.0, 5 mM MgCl<sub>2</sub>, 0.3 mM EDTA, 5 μM Phospho-Glycogen synthase peptide-2 as substrate, GSK-3β (0.1 μg/mL), 10 μM ATP ([γ-<sup>33</sup>P]-ATP; 2 μCi/mL), 0.001% Brij35, 0.01% 2-mercaptoethanol, 0.5% glycerol, and BSA (0.1 mg/mL) in a final volume of 50 μL. After 60 min incubation at 25 °C, reaction was terminated by adding 100 μL of 100 mM phosphoric acid. All reaction solution was transferred to Unifilter plate (Whatman; UniFilter P81) as a filter plate using a harvester (PerkinElmer; FILTERMATE 196). After harvesting, the Unifilter plates were washed six times with 100 mM phosphoric acid and dried. These plates were counted in Topcount (PerkinElmer) by adding 20 μL of Microscint-O (PerkinElmer).

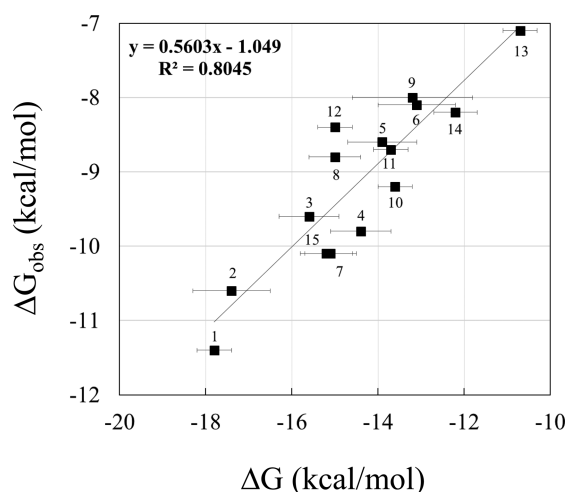
IC<sub>50</sub> measurement was done twice in a separate experiment. IC<sub>50</sub> values were determined in duplicates by constructing a dose–response curve at 10 different test compound concentrations (100, 50, 25, 12.5, 6.25, 3.12, 1.56, 0.78, 0.39, and 0.19 μM), using a four-parameter logistical model with XLfit software (IDBS, Emeryville, CA, USA).

## RESULTS AND DISCUSSION

**Relationship between Inhibitory Activity and Calculated ΔG.** Table 1 shows 15 chemical structures of 7-azaindole derivatives used in this study, together with their inhibitory activities (IC<sub>50</sub>) for GSK-3β. Calculated values of ΔG, ΔG<sub>sol</sub>, and ΔG<sub>complex</sub> and observed values of ΔG<sub>obs</sub> [= RT ln (IC<sub>50</sub>)] are also



**Figure 1.** A stereoview of docking structures of derivative 1 and inhibitor AR-A014418 to the GSK-3 $\beta$  ATP binding site. Dotted lines express common hydrogen bonds between inhibitors and the GSK-3 $\beta$  ATP binding site. Both inhibitors are in the similar docking manner. Derivative 1 was selected as a representative from other derivatives.



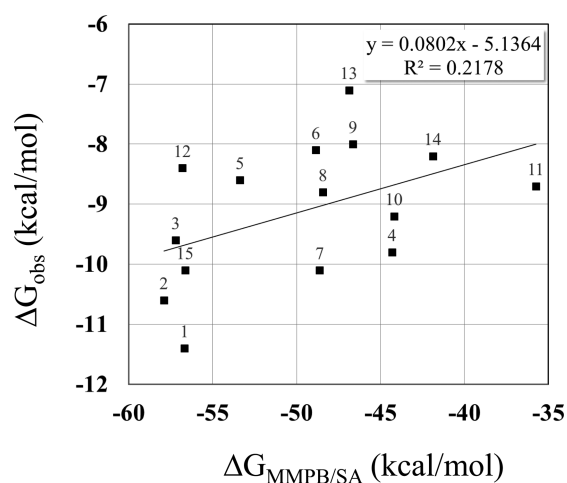
**Figure 2.** A scatter diagram of calculated and observed  $\Delta G$  values. Square of correlation coefficient ( $R^2$ ) and regression line are shown. Numbers correspond to derivatives in Table 1. Error bar mean deviation of calculated  $\Delta G$  as shown in Table 1.

given in Table 1. Figure 2 is a scatter diagram for calculated and observed  $\Delta G$ s.

It is remarkable to note that the diagram reveals a linear relation between both  $\Delta G$ s, where the correlation coefficient is 0.895 and the  $\Delta G$  value is a mere 1.6 times larger than the corresponding  $\Delta G_{\text{obs}}$ . This suggests that our 3D docking models for  $\Delta G$  calculation are almost acceptable. Concurrently, we could quantitatively predict inhibitory activities values of new 7-azaindole derivatives by using a regression line,  $Y = 0.55X - 1.13$ , where  $Y$  is the observed  $\Delta G_{\text{obs}}$ , i.e., inhibitory activity and  $X$  is the calculated one.

To know the utility of this  $\Delta G$  calculation on lead-optimization, we performed an additional 10 ns MD simulation on each GSK-3 $\beta$ /derivative complex molecule and calculated the “binding free energy” of each derivative by using MMPB/SA.

Figure 3 is a scatter diagram between  $\Delta G_{\text{obs}}$  values and those of MMPB/SA values for 15 derivatives. There is no high correlation between them, suggesting that MMPB/SA values do not sufficiently



**Figure 3.** A scatter diagram between MMPB/SA values and observed  $\Delta G$  values. MMPB/SA was applied to 100 sampling points in 10 ns MD simulation. The average value is shown in the figure.

estimate inhibitory activities of new 7-azaindole derivatives. Meanwhile, we also calculated MMGB/SA, in which the generalized Born model is used as the solvation energy. However, we could not observe any correlation between  $\Delta G_{\text{obs}}$  and the values of MMGB/SA as well (Figure S1).

We also calculated “docking ability” (i.e., score) by using other methods, that is, *LigScore*<sup>32</sup> and *PMF04*.<sup>33</sup> They have been frequently applied to *in silico* screening. The results, however, are no high correlations between  $\Delta G_{\text{obs}}$  values and scores calculated by these two methods: the correlation coefficients  $r = 0.119$  and  $r = 0.463$  for *LigScore* and *PMF04*, respectively, where final snapshots were applied to these two methods.

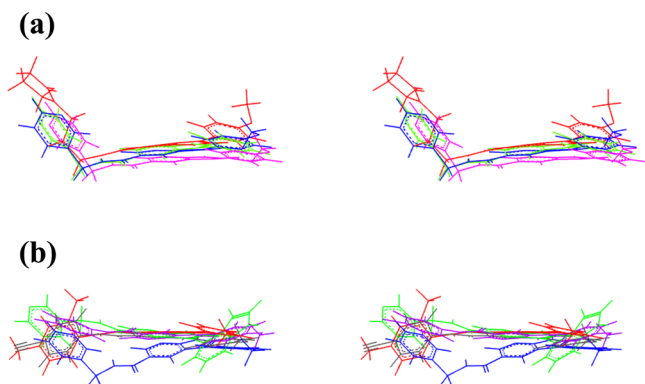
Although the undeniable results described here are not necessarily observed in the cases of other derivatives and target proteins, at least, this  $\Delta G$  calculation is very helpful to estimations of inhibitory activities of 7-azaindole derivatives.

**A Role of Hydrogen Bonding Formations to Backbone Amides.** We also tried to examine the MD trajectories of 15 complex molecules in order to know whether  $\Delta G_{\text{obs}}$  values



(or  $\Delta G$ ) are influenced by hydrogen bonding formations to the backbone amides. As a result, we found no relationship between them. For example, although 5 derivatives (**1**, **3**, **6**, **7**, and **11**) kept the same hydrogen bonding pattern to the backbone amides of Asp133 and Val135 during the 10 ns MD simulation, their  $IC_{50}$  and  $\Delta G$ s showed various values, as shown in Table 1. This suggests that hydrogen bonds to the backbone amides do not directly lead to high inhibitory activity on 7-azaindole derivatives. Hydrogen bonding formations to backbone amides rather play the basic role of helping 7-azaindole derivatives recognize the ATP binding site of GSK-3 $\beta$  or *vice versa*.

**Conformational Comparison between High and Low Active 7-Azaindole Derivatives.** Figure 4(a) shows a superimposition of 7-azaindole derivatives (**1**, **2**, **7**, and **15**) having high inhibitory activity. It is noticeable that their conformations are very similar each other. Figure 4(b) also



**Figure 4.** Stereoviews of superimposed conformations of 7-azaindole derivatives. Conformations were snapshots at 10 ns. This superimposition was based on 4 atoms, i.e., C=O carbonyl bonds of Asp133 and Val135 mainchain. (a) is a superimposition of derivatives with high activity. Derivative **1** is shown in blue, **2** in red, **7** in green, and **15** in magenta. (b) is a superimposition of low active derivatives and high active derivative **1**. Derivative **6** is shown in red, **9** in green, **13** in magenta, **14** in gray, and **1** in blue.

shows a superimposition of low active derivatives (**6**, **9**, **31**, and **14**) and derivative **1** as a representative of high active ones. The conformations of low active derivatives are also similar to each other and are significantly different from that of derivative **1**. Thus, we supposed that inhibitory activity depends on conformational differences when 7-azaindole derivatives dock into the ATP binding site of GSK-3 $\beta$ . Such a conformational difference is mainly due to the sum of van der Waals forces among a large number of atoms belonging to derivatives, the ATP binding site, and neighboring water molecules. Indeed, we could not find any functional groups contributing to inhibitory activity.

**Thermodynamic Behavior of Derivatives.** The  $\Delta G$  calculation serves to estimate the thermodynamic behavior of each derivative with GSK-3 $\beta$ . To show this behavior revealed by  $\Delta G$  calculation, we focus herein on derivatives **7** and **15**, which exhibit the same inhibitory activities (see Table 1). Comparison between their  $\Delta G_{\text{complex}}$  values suggests that **15** binds to the ATP binding site more easily than does **7** since the absolute value of the  $\Delta G_{\text{complex}}$  of derivative **15** is larger than that of **7**. Simultaneously, their  $\Delta G_{\text{sol}}$  values also suggest that **15** could be distributed into the GSK-3 $\beta$  side less than could **7** because **15** is more hydrophilic than is **7**. Conversely, **7** could distribute to the GSK-3 $\beta$  side more than could **15**, although the binding ability of

**7** to the ATP binding site is rather less than that of **15**. Namely, these two derivatives are not equivalent to each other for their behavior from the standpoint of thermodynamics, although their  $\Delta G$ s are nearly equal. Similar thermodynamic behaviors could be observed for the **5** and **11**, **6** and **9**, **3** and **4**, and **4** and **7** pairs. Such thermodynamic behaviors suggest that even if the molecular structure of an inhibitor fits well to the ATP binding site, it does not directly link with the expected high activity in the case of hydrophilic derivatives; in other words, even if a binding structure with the ATP binding site is to some extent loose, the derivative exhibits high activity when it is hydrophobic. Thus, it could be concluded that inhibitor activity with the target protein is influenced not only by structural complementarity between their molecules but also by the ratio of inhibitor distributed between the water side and target protein side.

In actual lead-optimization, we have frequently experienced that inhibitors designed with high structural complementarity do not always exhibit as high activity as could be expected. This has been a mystery in the drug design field. As mentioned above, however, we believe that the mystery is caused by optimization without considering the thermodynamic behavior of the designed inhibitor. Because of adhering to the “lock and key” concept, the present tactics for inhibitor design are too much predisposed to improving structural complementarity between the inhibitor and its target protein active site and so provide a platform for effects such as hydrogen bond formation, van der Waals and electrostatic interactions, and so on.

Here, we would like to emphasize that lead-optimization should be performed under consideration of thermodynamic behavior, especially for the distribution to the water side, i.e., solubility, together with structural complementarity to the target protein active site. Then, if we could simultaneously estimate the degrees of the inhibitor for its solubility and complex formation, we could decrease the amount of extra effort, i.e., trials and errors in the lead-optimization cycle (drug design  $\rightarrow$  synthesis  $\rightarrow$  measurements of activity and solubility  $\rightarrow$  drug design  $\rightarrow$  ...).

**A Proposal of Lead-Optimization Procedure Using  $\Delta G$  Calculation.** Along the line of efficient lead-optimization, we propose an optimization procedure based on binding free energy ( $\Delta G$ ) as follows.

first step: Finding several lead compounds from HTS, the crystal structure of compound-target protein complexes, or other information

second step: Determining scaffolds should be optimized on the basis of lead compounds

third step: Designing several derivatives on the basis of a typical scaffold and constructing their 3D docking models with the target protein and synthesizing their derivatives

fourth step: Calculating  $\Delta G$  and also measuring activity ( $IC_{50}$ ) for designed derivatives

fifth step: Checking the relationship between calculated and observed  $\Delta G$ s. In the case of no significant correlation between them, then reconstructing the 3D docking model of designed derivatives

sixth step: After confirming a significant correlation, choosing derivatives with a relatively larger  $|\Delta G_{\text{sol}}|$  among derivatives showing a large  $|\Delta G|$  value

seventh step: Considering their chosen derivatives as new lead compounds and designing new derivatives, and calculating  $\Delta G$ s of their derivatives

eighth step: Choosing derivatives with large  $|\Delta G|$  values among calculated  $\Delta G$ s, then excluding derivatives with a smaller  $|\Delta G_{\text{sol}}|$

ninth step: Synthesizing derivatives with large  $|\Delta G|$  and measuring their activities ( $IC_{50}$ ) and solubilities

10th step: Repeating from the second to ninth steps for another scaffold if no promising compounds are optimized.

To certify the effectiveness of this process for inhibitor optimization based on activity and solubility, many more success cases should be accumulated. However, in our case on 7-azaindole derivatives, the number of compounds with both inhibitory activity and proper solubility was considerably sorted out from the number of candidates listed in the original synthetic plan, and we succeeded in bringing several candidates to the next stage of drug discovery, which carefully examines toxicities at the cell level and physicochemical properties required as a medical drug.

## CONCLUDING REMARKS

Binding between inhibitor and its target protein is controlled by the second law of thermodynamics. Thus, it is desirable to perform lead-optimization on the basis of binding free energy. As shown in this study, the calculation of binding free energy change enables the prediction of inhibitor activity. In addition, it provides information about the thermodynamic behavior of the inhibitor necessary for drug design. For these reasons, we believe the free energy calculation approach will surely play an important role in the lead-optimization stage.

## ASSOCIATED CONTENT

### Supporting Information

Table S1 and Figures S1 and S2. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*Phone: +81 (0)48-721-1155 (ext. 3205). Fax: +81 (0)48-721-6718. E-mail: [kunihironov0846@view.ocn.ne.jp](mailto:kunihironov0846@view.ocn.ne.jp).

### Author Contributions

K.K. designed the study, overall project stratagem, and management and wrote the paper. Y.T. and S.N. performed the MD simulations and collected data. T.U. measured inhibitory activity and bioassay. K.O. calculated the MMPB/SA and MMGB/SA values. R.H. managed the calculation. H.C. wrote the paper partially.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The results of calculations were performed by using the RIKEN Integrated Cluster of Clusters (RICC). The authors would like to express special thanks to Emeritus Prof. Toshimasa Ishida, Osaka University of Pharmaceutical Sciences, for his many comments and would also like to thank Dr. Shigeru Okuyama, Taisho Pharmaceutical Co., Ltd., for the use of biological data on 7-azaindole derivatives.

## REFERENCES

- (1) Merz, K. M., Jr.; Ringe, D.; Reynolds, C. H. *Drug Design: Structure- and Ligand-Based Approaches*; Cambridge University Press: Cambridge, 2010.
- (2) Schneider, G. *De novo Molecular Design*; Wiley-VCH Verlag GmbH & Co: Weinheim, 2014.
- (3) Kirkwood, J. G. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **1935**, *3*, 300–313.

- (4) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
- (5) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (6) Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water. *J. Chem. Phys.* **1988**, *89*, 3742–3746.
- (7) Pranata, J.; Jorgensen, L. W. Monte Carlo simulations yield absolute free energies of binding for guanine–cytosine and adenine–uracil base pairs in chloroform. *Tetrahedron* **1991**, *47*, 2491–2501.
- (8) Cohen, P.; Yellowlees, D.; Aitken, A.; Donella-Deana, A.; Hemmings, B. A.; Parker, P. J. Separation and Characterisation of Glycogen Synthase Kinase 3, Glycogen Synthase Kinase 4 and Glycogen Synthase Kinase 5 from Rabbit Skeletal Muscle. *Eur. J. Biochem.* **1982**, *124*, 21–35.
- (9) Summers, S. A.; Kao, A. W.; Kohn, A. D.; Backus, G. S.; Roth, R. A.; Pessin, J. E.; Birnbaum, M. J. The role of glycogen synthase kinase 3 $\beta$  in insulin-stimulated glucose metabolism. *J. Biol. Chem.* **1999**, *274*, 17934–17940.
- (10) Li, B.; Ryder, J.; Su, Y.; Moore, S. A., Jr.; Liu, F.; Solenberg, P.; Brune, K.; Fox, N.; Ni, B.; Liu, R.; Zhou, Y. Overexpression of GSK3 $\beta$  S9A resulted in tau hyperphosphorylation and morphology reminiscent of pretangle-like neurons in the brain of PDGSK3 $\beta$  transgenic mice. *Transgenic Res.* **2004**, *13*, 385–396.
- (11) Stoothoff, W. H.; Johnson, G. V. Tau phosphorylation: physiological and pathological consequences. *Biochim. Biophys. Acta* **2005**, *1739*, 280–297.
- (12) Martinez, A. Preclinical efficacy on GSK-3 inhibitors: towards a future generation of powerful drugs. *Med. Res. Rev.* **2008**, *28*, 773–796.
- (13) Arnost, M.; Pierce, A.; ter Haar, E.; Lauffer, D.; Madden, J.; Tanner, K.; Green, J. 3-Aryl-4-(arylhydrazono)-1H-pyrazol-5-ones: Highly ligand efficient and potent inhibitors of GSK3 $\beta$ . *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1661–1664.
- (14) Coffman, K.; Brodney, M.; Cook, J.; Lanyon, L.; Pandit, J.; Sakya, S.; Schachter, J.; Tseng-Lovering, E.; Wessel, M. 6-Amino-4-(pyrimidin-4-yl)pyridones: novel glycogen synthase kinase-3 $\beta$  inhibitors. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 1429–1433.
- (15) Bhat, R.; Xue, Y.; Berg, S.; Hellberg, S.; Ormo, M.; Nilsson, Y.; Radesater, A. C.; Jerning, E.; Markgren, P. O.; Borgegard, T.; Nylof, M.; Gimenez-Cassina, A.; Hernandez, F.; Lucas, J. J.; Diaz-Nido, J.; Avila, J. Structural insights and biological effects of glycogen synthase kinase 3-specific inhibitor AR-A014418. *J. Biol. Chem.* **2003**, *278*, 45937–45945.
- (16) Cohen, P.; Goedert, M. GSK3 inhibitors: development and therapeutic potential. *Nat. Rev. Drug Discovery* **2004**, *3*, 479–487.
- (17) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (18) Donald, A.; McHardy, T.; Rowlands, M. G.; Hunter, L.-J. K.; Davies, T. G.; Berdini, V.; Boyle, R. G.; Aherne, G. W.; Garrett, M. D.; Collins, I. Rapid Evolution of 6-Phenylpurine Inhibitors of Protein Kinase B through Structure-Based Design. *J. Med. Chem.* **2007**, *50*, 2289–2292.
- (19) *Molecular Operating Environment (MOE)*, 2008.10; Chemical Computing Group, Inc. (CCG): Montreal, Quebec, Canada, 2008.
- (20) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (21) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An  $N \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (22) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–725.
- (23) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(24) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(25) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merck, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *AMBER 9*; AMBER: San Francisco, CA, 2006.

(26) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Rob, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.

(27) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(28) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(29) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Phys. Rev. Lett.* **2003**, *91*, 140601–140604.

(30) Fujitani, H.; Tanida, Y.; Ito, M.; Jayachandran, G.; Snow, C. D.; Shirts, M. R.; Sorin, E. J.; Pande, V. S. Direct calculation of the binding free energies of FKBP ligands. *J. Chem. Phys.* **2005**, *123*, 084108.

(31) Fujitani, H.; Tanida, Y.; Matsuura, A. Massively parallel computation of absolute binding free energy with well-equilibrated states. *Phys. Rev. E* **2009**, *79*, 021914.

(32) Krammer, A.; Kirchhoff, P. D.; Jiang, X.; Venkatachalam, C. M.; Waldman, M. LigScore: a novel scoring function for predicting binding affinities. *J. Mol. Graphics Modell.* **2005**, *23*, 395–407.

(33) Muegge, I. PMF Scoring Revisited. *J. Med. Chem.* **2005**, *49*, 5895–5902.

(34) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate–DNA Helices. *J. Am. Chem. Soc.* **1998**, *120*, 9401–9409.