

# Virtual-Ligand-Assisted Screening Strategy to Discover Enabling Ligands for Transition Metal Catalysis

Wataru Matsuoka,<sup>\*,†</sup> Yu Harabuchi,<sup>†,‡,§</sup> and Satoshi Maeda<sup>\*,†,‡,§,Δ</sup>

<sup>†</sup>Department of Chemistry, Graduate School of Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan

<sup>‡</sup>Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Sapporo, Hokkaido 001-0021, Japan

<sup>§</sup>ERATO Maeda Artificial Intelligence for Chemical Reaction Design and Discovery Project, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan

<sup>Δ</sup>Research and Services Division of Materials Data and Integrated System (MaDIS), National Institute for Materials Science (NIMS), Tsukuba, Ibaraki 305-0044, Japan

**ABSTRACT:** The ligand screening process, in which an optimal ligand for a reaction of interest is identified from an enormous and diverse set of candidate molecules, is of particular importance in the development of transition metal catalysis. Conventionally, this process has been performed by experimental trial-and-error cycles, which require significant time and resources. Herein, we report a novel strategy called “virtual-ligand-assisted (VLA) screening” that enables practical *in silico* ligand screening based on the transition state theory. We developed a virtual ligand,  $\text{PCl}^*_3$ , which parameterizes both the electronic and steric effects of monodentate phosphorus(III) ligands in quantum chemical calculations, and used it to assess how these effects perturb the energy profile of a reaction. This parameter-based ligand screening approach allowed us to identify the optimal electronic and steric effects for a reaction of interest, thereby affording guiding principles for rational ligand design. The VLA screening strategy was demonstrated for the selectivity-determining step of the rhodium-catalyzed hydroformylation of a terminal olefin, and phosphorus(III) ligands with potentially high linear or branched selectivities were designed. These findings indicate that VLA screening is a promising approach for streamlining the ligand screening process.

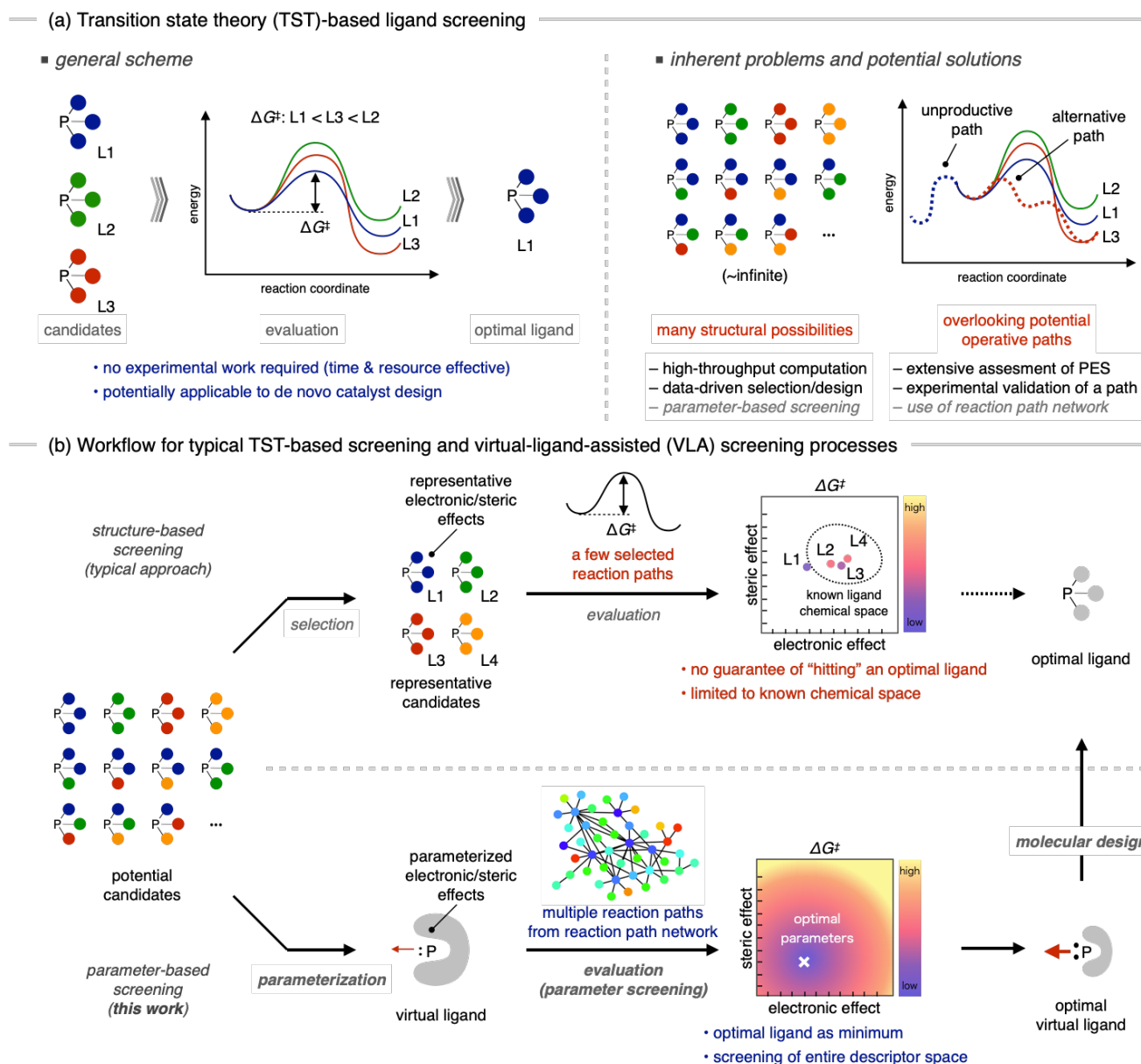
## 1. INTRODUCTION

Over the past several decades, the development of transition-metal-catalyzed reactions, such as hydroformylation,<sup>1</sup> olefin metathesis,<sup>2</sup> asymmetric hydrogenation,<sup>3</sup> and cross-coupling reactions,<sup>4</sup> has become a central research topic in synthetic chemistry. These advances in transition metal catalysis have significantly contributed to progress in broad scientific and technological fields. Considerable attention has been paid to the design of superior catalysts, including introducing organic ligands such as organophosphorus compounds and *N*-heterocyclic carbenes (NHCs), which often enhance catalytic activity and selectivity.<sup>1–5</sup> The effects of such ligands on metal center reactivity have been well studied, and it is widely accepted that both the electronic and steric aspects of a ligand change the reactivity of a metal center.<sup>6–8</sup> Because the extent of perturbation provided by a ligand (electronic and steric effects) is intrinsically determined by its chemical structures, ligand screening (or ligand design), in which many candidate molecules are examined to find an optimal ligand for the reaction of interest, is recognized as one of the most important processes in the development of transition metal catalysis.

Conventionally, this process has been conducted through experimental trial-and-error cycles. A ligand candidate is designed, prepared, and evaluated in the reaction of interest. Then, based on the experimental results, the

next candidate is designed. In many cases, bottlenecks occur in the preparation and evaluation of individual ligand candidates. Therefore, to streamline this process, chemical libraries for representative classes of ligands and high-throughput screening techniques have been developed.<sup>9</sup> More recently, many attempts have been made to accelerate the screening process using informatics techniques.<sup>10</sup> In this approach, experimental results are stored as a data set, and the relationship between ligand chemical structure and reaction performance is investigated. Regression analysis, which has been adopted in many such studies, allows chemists to rapidly capture important but nonintuitive ligand features, thus minimizing the number of experimental trials required to find optimal ligands.

An *ab initio* approach based on transition state theory (TST) has also been intensively studied (Figure 1a).<sup>11–14</sup> In this approach, the performance of each ligand candidate is evaluated based on quantum chemical calculations, and the optimal ligand is identified as the candidate that results in the best energy profile for the reaction (e.g., a low activation barrier). TST-based ligand screening can be theoretically conducted *in silico* without any experimental work. Thus, this approach is not only time and resource efficient but may also enable *de novo* catalyst design, which is often regarded as a “holy grail” in computational chemistry.<sup>11</sup>



**Figure 1.** (a) Schematic illustration of TST-based ligand screening and associated inherent problems. (b) Workflow for typical TST-based ligand screening and VLA screening processes.

However, there are still limited reports on ligand screening using the TST-based approach,<sup>11–14</sup> probably for the following reasons. First, the structural possibilities for ligands are almost infinite (Figure 1a, right). In general, because a transition metal complex has many conformational possibilities, numerous attempts and huge amounts of computational resources are required to determine the energy profile of a reaction, even for a single ligand candidate. In addition, considering the enormous variety of potential ligand structures, it is unrealistic to evaluate all candidates individually. Therefore, screening must be conducted by selecting a limited number of ligands as representative candidates (Figure 1b, top). Such heuristic choices not only cause truly optimal ligands to be overlooked but also narrow the chemical space to be explored, as known and readily available ligands tend to be unconsciously selected. Recently, to address this issue, high-

throughput computational techniques<sup>13</sup> and data-driven candidate selection<sup>14</sup> have been intensively investigated.

The second obstacle in TST-based ligand screening is the complexity of the reaction mechanism (Figure 1a, right). Because almost all calculations for transition metal catalysis are currently performed based on mechanistic assumptions by chemists, the obtained results are potentially biased or incorrect. Therefore, to ensure the reliability of the results, TST-based screening requires either extensive assessment of the potential energy surface (PES) by chemists or experimental validation of the obtained mechanism. This process not only decreases the practicality of TST-based ligand screening but may also overlook both nonintuitive productive and unproductive paths that operate when the ligand structure is modified. In theory, this problem can be (at least partially) avoided by performing ligand screening based on a reaction path network generated

by systematic automated reaction path search methods<sup>15,16</sup> instead of a few reaction paths proposed by chemists or validated in specific cases. However, these algorithms are often too computationally demanding for practical use, especially in the case of transition metal catalysis, where the large number of atoms in an organic ligand must be considered.

Considering this background, we herein report an efficient quantum chemical calculation method, namely, “virtual-ligand-assisted (VLA) screening”, to realize practical TST-based ligand screening (Figure 1b, bottom). A virtual ligand is a “dummy” ligand for quantum chemical calculations that reproduces the reactivity (e.g., the activation barrier of a reaction) of the corresponding transition metal complex by imitating the electronic and/or steric effects associated with the real ligand. In a pioneering work, Sakaki and coworkers developed PC<sup>#(R)</sup><sub>3</sub>, a virtual ligand that represents the electronic effect of the trialkylphosphine ligand (PR<sub>3</sub>) by tuning the quantum capping potential (C<sup>#(R)</sup>) to reproduce the lone pair orbital energy of PR<sub>3</sub>.<sup>17</sup> As PC<sup>#(R)</sup><sub>3</sub> treats the ligand substituents (R) implicitly, energy calculations for the corresponding transition metal complex can be performed rapidly, regardless of the approximated ligand. Inspired by this work, we anticipated that developing a virtual ligand that reproduces both the electronic and steric effects of a ligand<sup>6-8</sup> would dramatically reduce the computational costs not only for energy calculations but also for geometry optimization, including transition state determination. Such a virtual ligand would allow the application of automated reaction path search methods<sup>15,16</sup> to transition metal catalysis, thereby enabling TST-based ligand screening to be performed based on an unbiased reaction path network that includes both productive and unproductive paths. Moreover, by using the electronic and steric effects of the virtual ligand as parameters and tuning them to optimize the energy profile of the reaction rather than to reproduce a specific ligand, ligand screening can be realized virtually without heuristically selecting a few representative candidates (Figure 1b, bottom). It should be noted that the screening process is conducted based on continuous electronic and steric parameters instead of discontinuous chemical structures. This parameterization enables the determination of the optimal electronic and steric effects of ligands at the minimum point of a function (e.g., activation barrier) of these parameters, thus providing guiding principles for the molecular design of optimal (real) ligands. Furthermore, VLA screening allows exploration of the entire space determined by the electronic and steric parameters without any constraints originating from real ligands, such as synthetic accessibility or stability. Therefore, it is possible to find novel ligands both within and outside the known chemical space of ligands. Here, as a proof of concept, we developed a virtual ligand for one of the most common ligand classes, monodentate phosphorus(III) compounds, and performed VLA screening for the regioselective hydroformylation of terminal olefins.

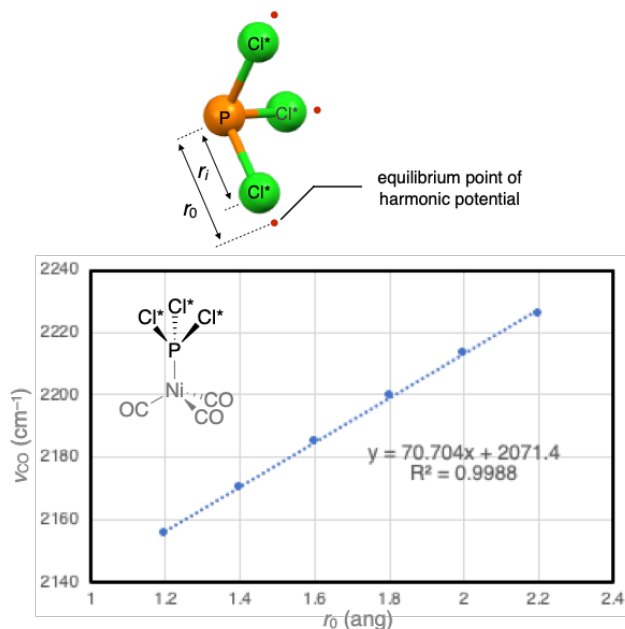
## 2. THEORY AND METHOD

### 2.1. Approximation of Electronic Effects

First, a newly developed method for approximating the electronic effects of phosphorus(III) ligands is described. Although many descriptors for the electronic effects of phosphorus(III) ligands have been proposed,<sup>6,7</sup> we focus on Tolman’s electronic parameter (TEP).<sup>7a</sup> The TEP of a ligand (L), which is defined by the frequency of the A<sub>1</sub> C–O vibrational mode ( $\nu_{CO}$ ) of corresponding (pseudo-)C<sub>3v</sub> symmetric complex LNi(CO)<sub>3</sub>, can be determined both experimentally and computationally.<sup>7b</sup> The TEP captures the net electronic effect (i.e., donation and back-donation) of a ligand,<sup>7b</sup> whereas the lone pair orbital energy, which was used as the electronic descriptor in Sakaki’s virtual ligand,<sup>17</sup> reflects only the electron-donating ability. Therefore, we assumed that the electronic effects of a phosphorus(III) ligand can be approximated well in quantum chemical calculations by reproducing the calculated TEP with a virtual ligand. Extensive attempts to develop such a system revealed that a wide range of TEP values can be reproduced using a trichlorophosphine-based virtual ligand, PCl<sub>3</sub><sup>\*</sup> (Figure 2, see Figure S1 for details). In this virtual ligand, the Cl atoms are placed under the harmonic potential (keep potential):

$$V_{keep} = \sum_{i=1}^3 \frac{1}{2} k(r_i - r_0)^2$$

where  $k$ ,  $r_i$ , and  $r_0$  are the force constant, distance between the P and Cl atoms, and equilibrium distance of the harmonic potential, respectively. The  $k$  value was set to  $4.48 \times 10^3$  kcal/(mol Å<sup>2</sup>) (2.0 a.u.). We performed geometry optimization and harmonic vibrational frequency analyses of (Cl<sub>3</sub>P)Ni(CO)<sub>3</sub> complexes with  $r_0$  values of 1.2–2.2 Å. The TEP value ( $\nu_{CO}$ ) of PCl<sub>3</sub><sup>\*</sup> was found to be proportional to the  $r_0$  value in the  $\nu_{CO}$  range of 2156–2226 cm<sup>−1</sup> (Figure 2). Notably, the calculated TEP values (at the same computational level) for highly electron-rich P<sup>t</sup>Bu<sub>3</sub> and electron-deficient PF<sub>3</sub> were 2175 and 2224 cm<sup>−1</sup>, respectively. These results suggest that PCl<sub>3</sub><sup>\*</sup> can reproduce the electronic effects of a broad range of phosphorus(III) ligands by appropriately tuning the  $r_0$  value.



**Figure 2.** Approximation of the electronic effects of phosphorus(III) ligands using newly developed virtual ligand  $\text{PCl}^*_3$ .

## 2.2. Approximation of Steric Effects

In this section, a newly developed method for approximating the steric effects of phosphorus(III) ligands is described. Among the various descriptors for the steric effects of phosphorus(III) ligands,<sup>6,8</sup> we focused on the cone angle, which was also investigated in Tolman's seminal work.<sup>6a</sup> The cone angle,  $\theta$ , of a symmetric phosphorus(III) ligand corresponds to the apex angle of a cone that originates from a metal center located 2.28 Å from the phosphorus atom and just touches the van der Waals surface of the outermost ligand atoms. It is widely accepted that the steric effect of a ligand, or the degree of steric repulsion between the ligand and substrates around the metal center, correlates well with the ligand cone angle.<sup>6a,b</sup> Therefore, to reproduce the steric effects of phosphorus(III) ligands with a virtual ligand, we designed a cone-shaped repulsive potential (cone potential) (Figure 3a). First, based on the cone angle definition,<sup>6a</sup> we constructed a cone with apex angle  $\theta$  located 2.28 Å from the phosphorus atom of  $\text{PCl}^*_3$ . Then, the steric repulsion between the ligand and the substrates was reproduced by placing every atom in substrate S (i.e., all atoms except for the metal and those in  $\text{PCl}^*_3$ ) under a repulsive potential, as follows:

$$V_{\text{cone}} = \sum_{i \in S} 4\varepsilon \left( \frac{\sigma}{r_i + a\sigma} \right)^{12}$$

$$r_i = \begin{cases} d_i \sin\left(\phi_i - \frac{\theta}{2}\right) & \left(\phi_i - \frac{\theta}{2} \leq \frac{\pi}{2}\right) \\ d_i & \left(\phi_i - \frac{\theta}{2} > \frac{\pi}{2}\right) \end{cases}$$

where  $r_i$  is the distance between atom  $i$  (in S) and the cone surface;  $d_i$  is the distance between atom  $i$  and the cone apex;  $\phi_i$  is the angle defined by the phosphorus atom, the apex, and atom  $i$ ; the potential function,  $V_{\text{cone}}$ , is determined with reference to the repulsive term of the Lennard-Jones potential; and constants  $\varepsilon$ ,  $\sigma$ , and  $a$  are determined using another steric descriptor,  $\text{He}_8\text{steric}$  (Figure 3b).<sup>6b</sup>  $\text{He}_8\text{steric}$ , which was proposed by Harvey and coworkers, is defined by the interaction energy between the phosphorus(III) ligand and a ring of eight helium atoms. The helium atoms are held in regular positions on a circle of radius 2.5 Å, whereas the phosphorus atom in the ligand is fixed 2.28 Å from the center of the ring. Harvey and coworkers revealed that the  $\text{He}_8\text{steric}$  value for a phosphorus(III) ligand ( $E_{\text{ster}}$ ) is correlated with its cone angle ( $\theta$ ), as shown in Figure 3b (blue dots).<sup>6b</sup> Using the definitions of  $\text{He}_8\text{steric}$  and cone potential, the  $E_{\text{ster}}$  value exerted by the cone potential of  $\text{PCl}^*_3$  with cone angle  $\theta$  can be numerically calculated as follows:

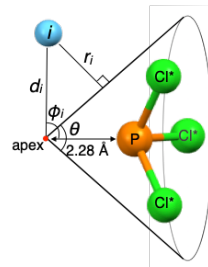
$$E_{\text{ster}} = 8 \cdot 4\varepsilon \left( \frac{\sigma}{r_{\text{He}} + a\sigma} \right)^{12}$$

$$r_{\text{He}} = 2.5 \sin\left(\frac{\pi - \theta}{2}\right)$$

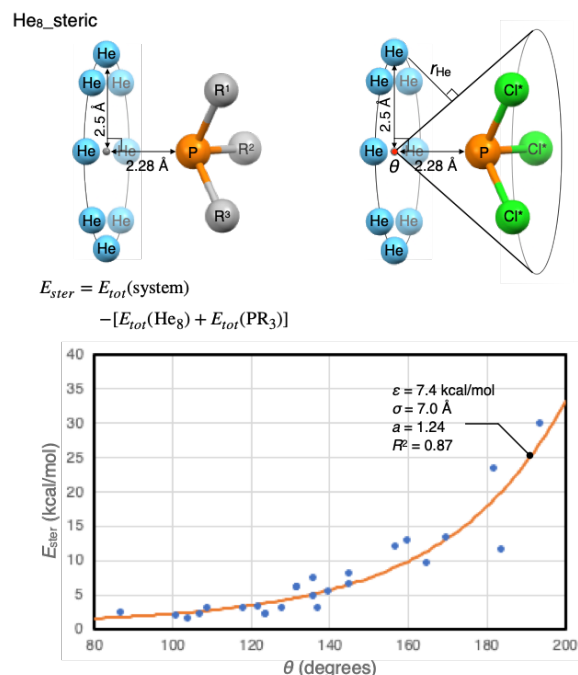
Constants  $\varepsilon$ ,  $\sigma$ , and  $a$  were optimized using a grid search so that the relationship between  $E_{\text{ster}}$  and  $\theta$  of the cone potential in the above equations reproduced the correlation between  $E_{\text{ster}}$  and  $\theta$  in real systems (blue dots, Figure 3b). The grid search was performed for  $\varepsilon$  values in the range of 1–

11 kcal/mol with a 0.1 kcal/mol step width,  $\sigma$  values in the range of 1–11 Å with a 0.1 Å step width, and  $a$  values in the range of 0–2 with a 0.01 step width. As a result, constants  $\varepsilon$ ,  $\sigma$ , and  $a$  were optimized as 7.4, 7.0, and 1.24, respectively, with a coefficient of determination ( $R^2$ ) of 0.87. The plot of  $E_{\text{ster}}$  against  $\theta$  for the cone potential with the optimal constants is shown in Figure 3b (orange line).

### (a) Definition of cone potential



### (b) Optimization of constants



**Figure 3.** (a) Cone potential definition for approximating steric effects. (b) Optimization of the cone potential function based on  $\text{He}_8\text{steric}$ .<sup>6b</sup>

## 2.3. Ligand-Ligand Repulsion

To accurately describe transition metal complexes bearing multiple phosphorus(III) ligands, the repulsion between a pair of virtual ligands  $\text{PCl}^*_3$  (i.e., ligand-ligand repulsion) must be evaluated. In this study, the repulsion between the virtual ligands was represented by the repulsion between two cones (Figure 4, top). However, this repulsion is slightly more complicated to implement than repulsion between a cone and an atom (Figure 3a) because the magnitude of ligand-ligand repulsion reflects not only the distance between the two ligands but also their relative configuration. Therefore, we treated the repulsion between two virtual ligands using the following simplified model (Figure 4, bottom). First, instead of repulsion be-

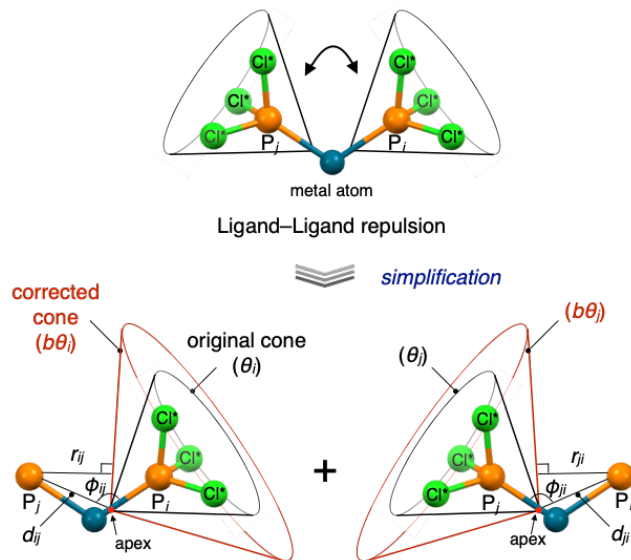


tween two virtual ligands, repulsion between one virtual ligand (e.g., a ligand containing a  $P_i$  atom) and the phosphorus atom of another virtual ligand ( $P_j$ ) is evaluated. Because  $P_j$  is located inside the cone, which is subjected to a repulsive force, this evaluation will always underestimate the ligand–ligand repulsion. To correct for this systematic error, we defined a larger cone (red cone, Figure 4) that shares the apex and axis of the original cone, and the ligand–ligand repulsion was estimated based on the corrected cone, as follows:

$$V_{\text{lig-}i\text{g}} = \sum_{i \in P_{\text{virt}}} \sum_{j \in P_{\text{virt}}, i \neq j} \frac{1}{2} \cdot 4\epsilon \left( \frac{\sigma}{r_{ij} + a\sigma} \right)^{12}$$

$$r_{ij} = \begin{cases} d_{ij} \sin\left(\phi_{ij} - \frac{b\theta_i}{2}\right) & \left(\phi_{ij} - \frac{b\theta_i}{2} \leq \frac{\pi}{2}\right) \\ d_{ij} & \left(\phi_{ij} - \frac{b\theta_i}{2} > \frac{\pi}{2}\right) \end{cases}$$

where  $\epsilon$ ,  $\sigma$ , and  $a$  are the constants determined in Section 2.2;  $P_{\text{virt}}$  contains all the phosphorus atoms of the virtual ligands in the system;  $r_{ij}$  and  $d_{ij}$  are the distance from the phosphorus atom subjected to the repulsive force ( $P_j$ ) to the surface and apex of the corrected cone, respectively;  $\phi_{ij}$  is the angle defined by the two phosphorus atoms and the apex of the cone including  $P_i$ ; and  $b$  is a constant that determines the relative size of the corrected cone. To avoid doubly counting ligand–ligand repulsions (i.e., repulsion from  $P_i$  to  $P_j$  and from  $P_j$  to  $P_i$ ), the magnitude of the repulsive potential was halved. The constant  $b$  was set to 1.7 based on the optimized structure of  $\text{HRh}(\text{CO})(\text{ethylene})(\text{PR}_3)_2$  complexes (see Figure S2 for details).



**Figure 4.** Simplified model for approximating ligand–ligand repulsion using virtual ligands.

### 3. RESULTS AND DISCUSSION

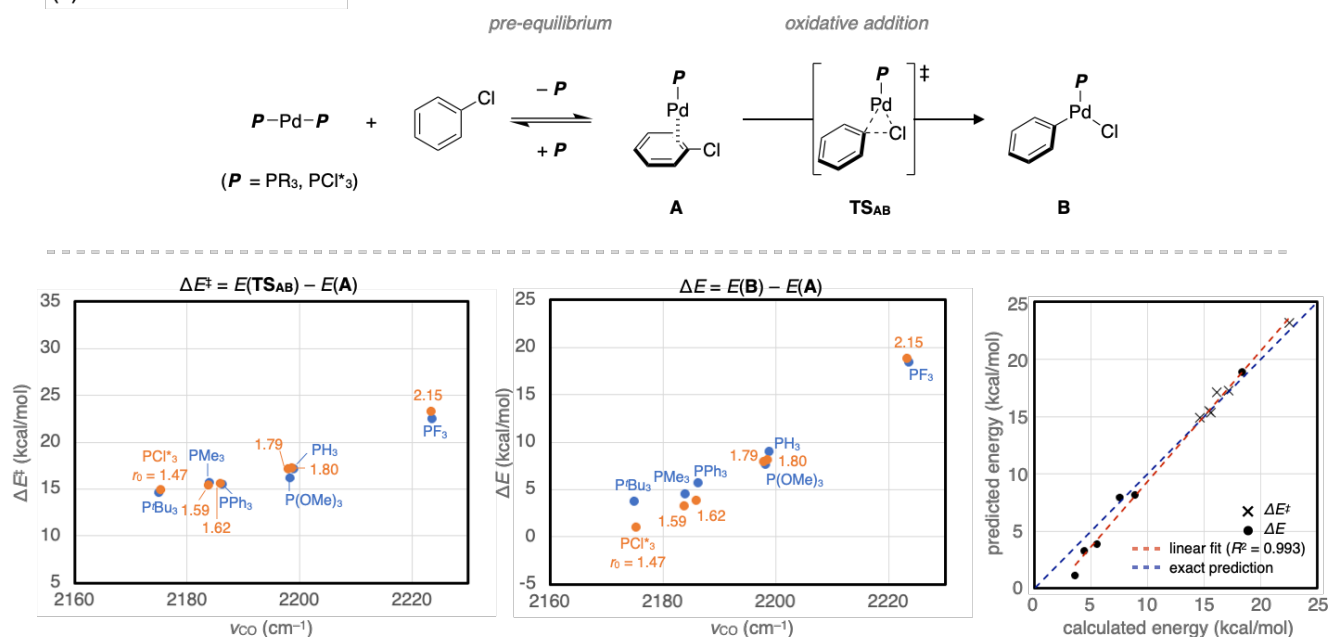
#### 3.1. Validation of Electronic Effect Approximation

The oxidative addition of chlorobenzene to a palladium(0) phosphine complex<sup>18</sup> was chosen as the target system to validate the electronic effect approximation for phosphorus(III) ligands using virtual ligand  $\text{PCl}_3^*$  (Figure

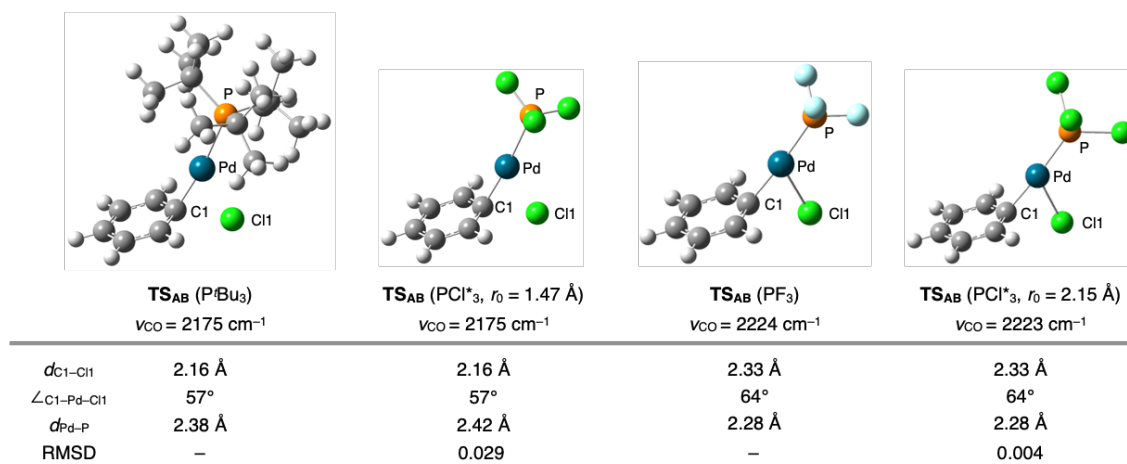
5). The oxidative addition of aryl chloride is known to proceed through three-membered transition state  $\text{TS}_{\text{AB}}$  after the formation of prereactive complex **A** (Figure 5a, top).<sup>18b–d</sup> We optimized the geometries of intermediate **A**, transition state  $\text{TS}_{\text{AB}}$ , and product **B** with six different phosphorus(III) ligands ( $\text{P}^t\text{Bu}_3$ ,  $\text{PMe}_3$ ,  $\text{PPh}_3$ ,  $\text{PH}_3$ ,  $\text{P}(\text{OMe})_3$ , and  $\text{PF}_3$ ), and then calculated activation energy  $\Delta E^\ddagger$  (energy difference between **A** and  $\text{TS}_{\text{AB}}$ ) and reaction energy  $\Delta E$  (energy difference between **A** and **B**) for each real system. The obtained  $\Delta E^\ddagger$  and  $\Delta E$  values were plotted against the TEP values ( $v_{\text{CO}}$ ) of the ligands (blue dots, Figure 5a). Notably, despite large differences in cone angle, the ligands with similar  $v_{\text{CO}}$  values exhibited comparable  $\Delta E^\ddagger$  and  $\Delta E$  values (compare  $\text{PH}_3$  with  $\text{P}(\text{OMe})_3$  or  $\text{PMe}_3$  with  $\text{PPh}_3$ ), which implies that electronic effects are of significance for this reaction. A series of calculations were also performed using six virtual ligands, each of which was tuned to reproduce the electronic effect of a real ligand (orange dots, Figure 5a). Specifically, the  $r_0$  value for the keep potential of a virtual ligand was set to reproduce the TEP value ( $v_{\text{CO}}$ ) of a real ligand using the relationship between  $v_{\text{CO}}$  and  $r_0$  (Figure 2), whereas the steric effect approximation using the cone potential was not considered so that the “pure” electronic effects of the ligands could be discussed. The  $\Delta E^\ddagger$  and  $\Delta E$  values calculated using the virtual ligand showed good agreement with those determined for the real systems. As shown in Figure 5a (right), linear regression of the energies ( $\Delta E^\ddagger$  and  $\Delta E$ ) calculated for the virtual and real systems yielded a very high coefficient of determination ( $R^2 = 0.993$ ). In addition, the regression line (red dashed line) nearly coincided with the diagonal (blue dashed line), representing an exact prediction.

To our delight, in addition to the energies, the optimized geometries of both the ground and transition states were reproduced well using the virtual ligand. As representative examples, the optimized structures of  $\text{TS}_{\text{AB}}$  with electron-rich  $\text{P}^t\text{Bu}_3$  and electron-deficient  $\text{PF}_3$  are shown in Figure 5b ( $\text{TS}_{\text{AB}}(\text{P}^t\text{Bu}_3)$  and  $\text{TS}_{\text{AB}}(\text{PF}_3)$ , respectively). For comparison, the structures of  $\text{TS}_{\text{AB}}$  featuring  $\text{PCl}_3^*$  with  $r_0$  values set to reproduce the electronic effects of  $\text{P}^t\text{Bu}_3$  and  $\text{PF}_3$  are also shown ( $\text{TS}_{\text{AB}}(\text{PCl}_3^*, r_0 = 1.47 \text{ \AA})$  and  $\text{TS}_{\text{AB}}(\text{PCl}_3^*, r_0 = 2.15 \text{ \AA})$ , respectively). A comparison of  $\text{TS}_{\text{AB}}(\text{P}^t\text{Bu}_3)$  and  $\text{TS}_{\text{AB}}(\text{PF}_3)$  shows that these two structures differ largely in the distance between the C1 and Cl atoms ( $d_{\text{C1-Cl}}$ ), the angle defined by the C1, Pd, and Cl atoms ( $\angle_{\text{C1-Pd-Cl}}$ ), and the distance between the Pd and P atoms ( $d_{\text{Pd-P}}$ ). In contrast, comparisons of  $\text{TS}_{\text{AB}}(\text{PCl}_3^*, r_0 = 1.47 \text{ \AA})$  with  $\text{TS}_{\text{AB}}(\text{P}^t\text{Bu}_3)$  and  $\text{TS}_{\text{AB}}(\text{PCl}_3^*, r_0 = 2.15 \text{ \AA})$  with  $\text{TS}_{\text{AB}}(\text{PF}_3)$  show that the structures optimized with the virtual ligand are quite similar to those optimized with the corresponding real ligand. The root-mean-square deviations (RMSDs) of the atomic positions determined after removing the substituents on the phosphorus atom were 0.029 for  $\text{TS}_{\text{AB}}(\text{PCl}_3^*, r_0 = 1.47 \text{ \AA})$  and  $\text{TS}_{\text{AB}}(\text{P}^t\text{Bu}_3)$  and 0.004 for  $\text{TS}_{\text{AB}}(\text{PCl}_3^*, r_0 = 2.15 \text{ \AA})$  and  $\text{TS}_{\text{AB}}(\text{PF}_3)$ . These results, as well as those shown in Figure 5a, suggest that the electronic effects of phosphorus(III) ligands are successfully reproduced and parameterized by virtual ligand  $\text{PCl}_3^*$ . It should be noted that the computational costs for calculating the transition metal complexes were dramatically reduced because regardless of the number of ligand atoms (e.g., 40 atoms for  $\text{P}^t\text{Bu}_3$ ), the ligand could be replaced with only four atoms ( $\text{PCl}_3$ ).

### (a) Oxidative addition of PhCl



### (b) Structure comparison

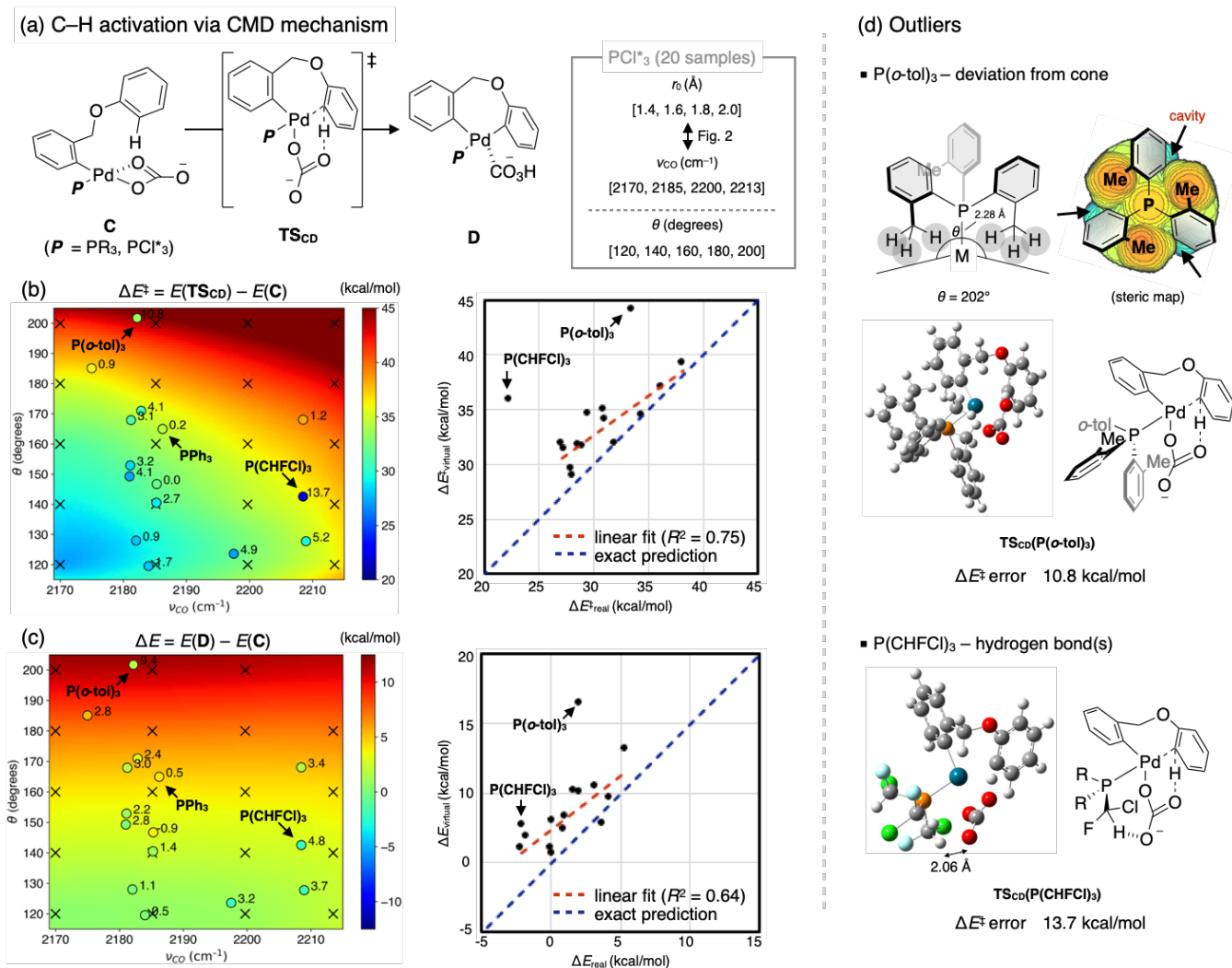


**Figure 5.** Validation of the electronic effect approximation with virtual ligand PCl\*<sub>3</sub> using the oxidative addition of aryl chloride to palladium(0) species as a model reaction.

### 3.2. Validation of Steric Effect Approximation

Intramolecular C–H activation through a concerted metalation-deprotonation (CMD) mechanism from arylpalladium **C** (Figure 6a)<sup>19</sup> was chosen as the model reaction to validate the approximation of steric effects with the virtual ligand. In contrast to the oxidative addition of chlorobenzene (Section 3.2), where the steric effect of the phosphorus(III) ligand was negligible, this reaction is expected to be sensitive to both electronic and steric effects because the palladium atom adopts a sterically crowded square-planar configuration. First, we optimized the geometries of **C**, **D**, and TS<sub>CD</sub> with 15 different real phosphorus(III) ligands (see Figures S3 and S4 for chemical structures), and then calculated activation energy ΔE<sup>‡</sup> (energy difference between **C** and TS<sub>CD</sub>) and reaction energy ΔE (energy difference between **C** and **D**) for each real system. These values were plotted against the TEP values (ν<sub>CO</sub>) and cone angles (θ) of the ligands (circles, Figures 6b and c, left). In

these plots (so-called Tolman's maps),<sup>6a</sup> the position of a circle represents the electronic and steric effects of a ligand, and its color corresponds to the calculated ΔE<sup>‡</sup> or ΔE value. Subsequently, the geometries of **C**, **D**, and TS<sub>CD</sub> were also optimized with virtual ligand PCl\*<sub>3</sub>. To evaluate the effects of electronic and steric parameters on ΔE<sup>‡</sup> or ΔE, 20 virtual systems were examined, with the r<sub>0</sub> value of the keep potential set to 1.4, 1.6, 1.8, or 2.0 Å and the θ value of the cone potential set to 120, 140, 160, 180, or 200°. The geometry optimization of complex **D**, which feature PCl\*<sub>3</sub> with r<sub>0</sub> and θ values of 2.0 Å and 200°, respectively, failed because of ligand dissociation. Hence, ΔE<sup>‡</sup> and ΔE were calculated for the remaining systems and plotted in Figures 6b and c (crosses) after converting the r<sub>0</sub> value of each virtual ligand to ν<sub>CO</sub> using the relationship in Figure 2. Then, the contour maps in Figures 6b and c were created by interpolating these data points. The ΔE<sup>‡</sup> and ΔE values for each real ligand were predicted using the contour maps, and the errors between these values (ΔE<sup>‡</sup><sub>virtual</sub>) and those



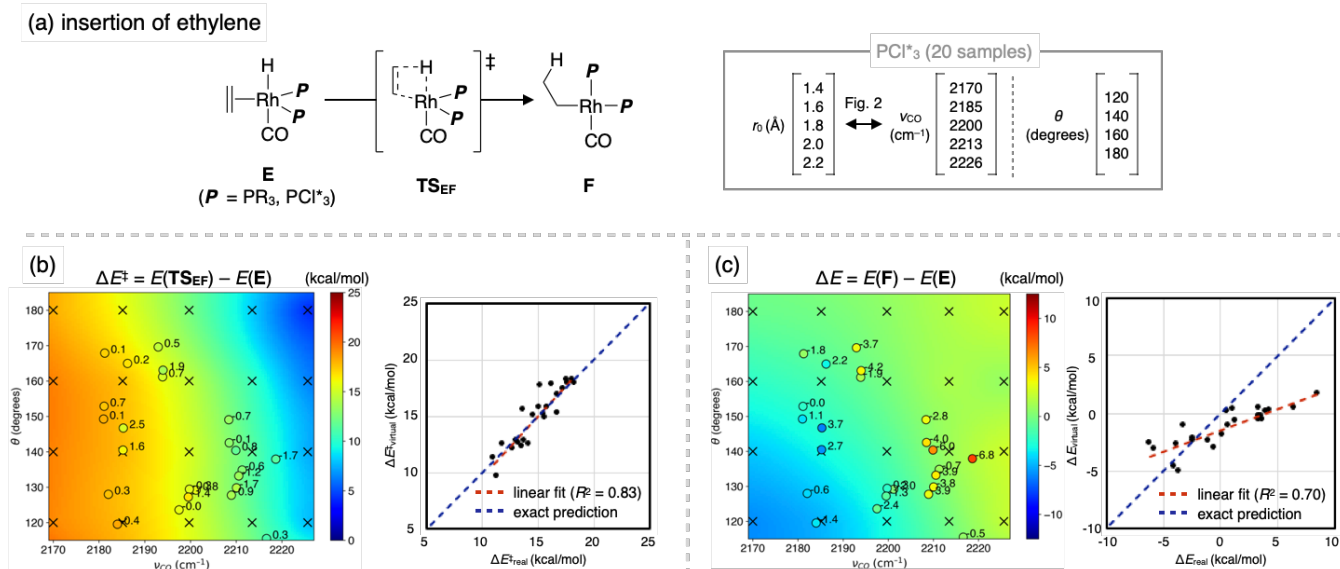
**Figure 6.** Validation of the steric effect approximation with virtual ligand  $\text{PCl}^*_3$  using intramolecular C–H activation at palladium(II) species as a model reaction.

calculated using the corresponding real systems ( $\Delta E^\ddagger_{\text{real}}$ ) are denoted at the top right of each circle. Finally, the  $\Delta E^\ddagger_{\text{virtual}}$  values were plotted against the  $\Delta E^\ddagger_{\text{real}}$  values to quantify the performance of the virtual ligand (Figures 6b and c, right).

As shown in Figure 6b (left), in both the real and virtual systems, ligands with a low TEP value and a small cone angle tended to have a small  $\Delta E^\ddagger$  value, except for two outliers among the real systems ( $\text{P}(\text{o-tol})_3$  and  $\text{P}(\text{CHFCl})_3$ , as discussed below in detail). Linear regression of  $\Delta E^\ddagger_{\text{virtual}}$  and  $\Delta E^\ddagger_{\text{real}}$  after removing the two outliers (Figure 6b, right) yielded a sufficient  $R^2$  value (0.75), although the regression line (red dashed line) deviated from the diagonal (blue dashed line) to some extent. Although the prediction of  $\Delta E^\ddagger$  by the virtual system is not quantitatively correct, these results imply that the virtual ligand captures the steric and electronic effects of the real ligands qualitatively. The systematic deviation of  $\Delta E^\ddagger_{\text{virtual}}$  from  $\Delta E^\ddagger_{\text{real}}$  suggests that the current virtual ligand model omits some important factors that exist in real ligands, probably because the steric effect approximation was overly simplified (see the below discussion for extreme examples). Thus, there is likely still scope for further improvement.

Similarly, the contour map for  $\Delta E$  qualitatively reproduced the trend in the real systems (Figure 6c). In both the real and virtual systems, the electronic effects of the ligands were found to be small, and ligands with a large cone angle tended to have a large  $\Delta E$  value. The regression line for  $\Delta E_{\text{virtual}}$  and  $\Delta E_{\text{real}}$  had an  $R^2$  value of 0.64, implying good correlation between these values.

Nevertheless, the predictions of  $\Delta E^\ddagger$  and  $\Delta E$  for  $\text{P}(\text{o-tol})_3$  and  $\text{P}(\text{CHFCl})_3$  were not successful, even qualitatively, for the following reasons. First, the cone angle seems to provide an inaccurate description of the steric effect for  $\text{P}(\text{o-tol})_3$ . As described in Section 2.2, virtual ligand  $\text{PCl}^*_3$  simulates the steric effect of the phosphorus(III) ligand using the cone potential obtained from the cone angle of the ligand. By definition,<sup>6a</sup> the cone angle of  $\text{P}(\text{o-tol})_3$  was determined by the hydrogen atoms of the methyl groups (Figure 6d, top). However, according to the steric map<sup>20</sup> of  $\text{P}(\text{o-tol})_3$ , cavities exist between the three methyl groups, which can be occupied by substrates or other ligands. Thus, the cone angle determined using the hydrogen atoms ( $202^\circ$ ) likely overestimates the steric effect of  $\text{P}(\text{o-tol})_3$ . Indeed, in the optimized structure of  $\text{TS}_{\text{CD}}$  with  $\text{P}(\text{o-tol})_3$ , the carbonate anion was found located in a cavity (Figure 6d,



**Figure 7.** Validation of steric effect approximation with virtual ligand PCI\*<sub>3</sub> using the insertion of ethylene from a rhodium(I) complex as a model reaction.

middle), which would cause the “effective” cone angle of P(*o*-tol)<sub>3</sub> to be much smaller than 202°. If the cone angle of P(*o*-tol)<sub>3</sub> is assumed to be similar to that of PPh<sub>3</sub> (165°), the predictions of  $\Delta E^\ddagger$  and  $\Delta E$  seem to be successful. In contrast, the large errors in the  $\Delta E^\ddagger$  and  $\Delta E$  value predicted for P(CHFCl)<sub>3</sub> seem to originate from a hydrogen bond between the ligand and a carbonate anion. In the optimized geometry of TS<sub>CD</sub> with P(CHFCl)<sub>3</sub>, the distance of 2.06 Å was observed between a hydrogen atom in the ligand and an oxygen atom in the carbonate anion (Figure 6d, bottom). This value is significantly smaller than the sum of the van der Waals radii of hydrogen and oxygen atoms (2.72 Å), indicating the existence of attractive interactions. Because the current virtual ligand model is not designed to reproduce attractive interactions such as hydrogen bonds, C–H– $\pi$  interactions, and  $\pi$ – $\pi$  interactions between ligands and substrates, the existence of such interactions causes large prediction errors.<sup>21</sup>

The insertion of ethylene from rhodium hydride complex **E** to give alkylrhodium complex **F** (Figure 7a)<sup>22</sup> was also investigated to further verify the approximation of electronic and steric effects with the virtual ligand. Because these rhodium complexes have two phosphorus(III) ligands, ligand–ligand repulsion between two virtual ligands (introduced in Section 2.3) needs to be considered along with ligand–substrate repulsion. Geometry optimization was performed and activation energy  $\Delta E^\ddagger$  and reaction energy  $\Delta E$  were evaluated for 24 phosphorus(III) ligands (real systems, see Figures S5 and S6 for chemical structures) as well as for PCI\*<sub>3</sub> (virtual systems) with 20 sets of electronic and steric parameters (Figures 7b and c). The trends for  $\Delta E^\ddagger$  and  $\Delta E$  calculated using the real systems were successfully reproduced by the virtual systems. In both the real and virtual systems,  $\Delta E^\ddagger$  tends to be small when an electron-deficient and bulky ligand is employed, whereas  $\Delta E$  tends to be small (exothermic reaction) when an electron-donating and less bulky ligand is employed. Gratifyingly, the prediction of  $\Delta E^\ddagger$  using the virtual ligand seemed to be quantitatively correct in this case. Specifical-

ly, the linear regression of  $\Delta E^\ddagger_{\text{virtual}}$  and  $\Delta E^\ddagger_{\text{real}}$  afforded a regression line ( $R^2 = 0.83$ ) that coincided with the diagonal, showing perfect correspondence (Figure 7b).

Considering the results shown in Figures 6 and 7, both the electronic and steric effects of phosphorus(III) ligands are, at least qualitatively, reproduced and parameterized by virtual ligand PCI\*<sub>3</sub> through the TEP and cone angle values.

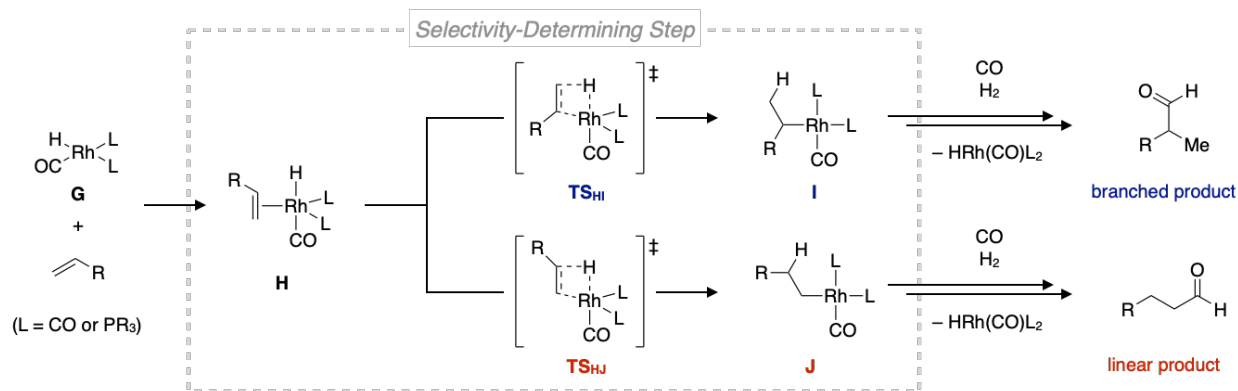
### 3.3. VLA Screening for the Regioselective Hydroformylation of Terminal Olefins

With the approximation of the electronic and steric effects of phosphorus(III) ligands by virtual ligand PCI\*<sub>3</sub> validated, we demonstrated the use of VLA screening to achieve the rhodium-catalyzed regioselective hydroformylation of terminal olefins (Figure 8a).<sup>1,22–24</sup> This reaction is considered to occur through the coordination of an olefin to rhodium complex **G** followed by the insertion of the olefin from complex **H**, as proposed by Wilkinson and coworkers.<sup>22a</sup> When a terminal olefin is employed as the substrate, two regioisomeric products (branched and linear products) can be formed. When olefin insertion proceeds via TS<sub>HI</sub> to give alkylrhodium complex **I**, the corresponding branched product is obtained. In contrast, when olefin insertion proceeds through TS<sub>HJ</sub>, the linear product is obtained. It is reasonable and widely accepted that the regioselectivity of the reaction is determined during the insertion of the olefin from complex **H**. Therefore, we assumed that the regioselectivity was kinetically determined by the difference in the Gibbs energies of TS<sub>HI</sub> and TS<sub>HJ</sub> (see Figures S7 and S8 for theoretical support) and conducted VLA screening to design phosphorus(III) ligands that realize the regioselective hydroformylation of terminal olefins.

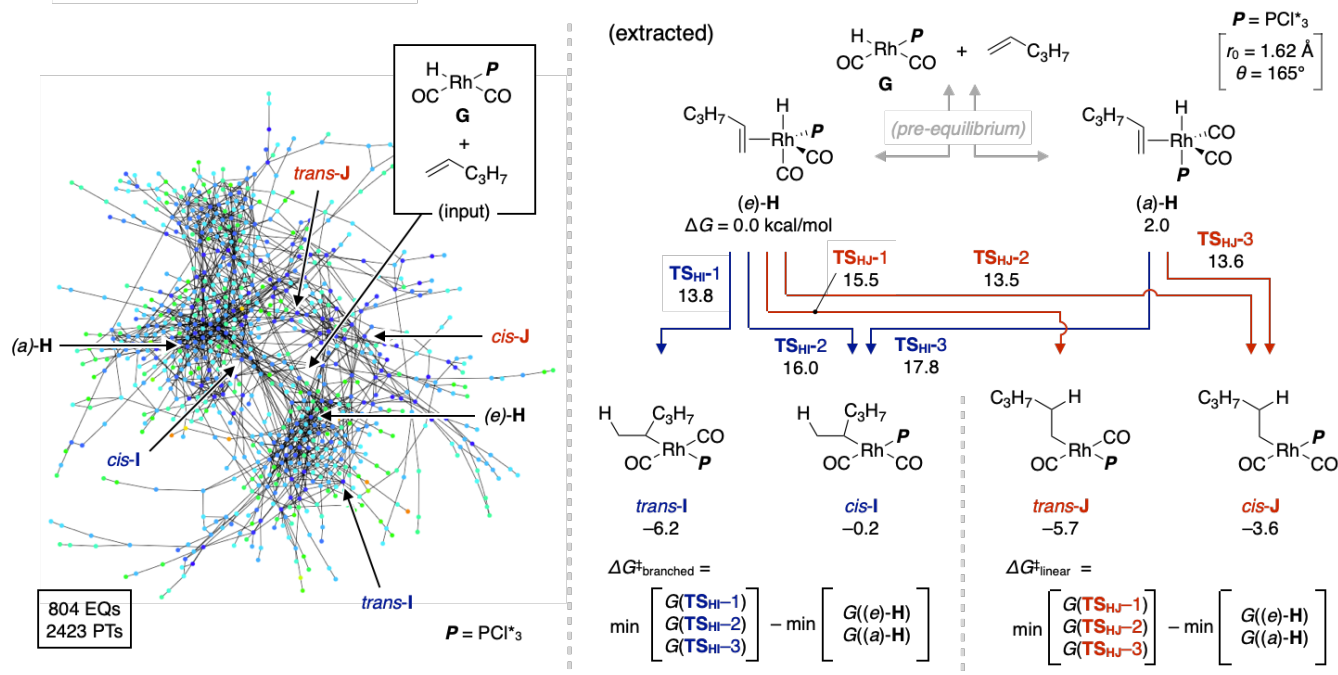
As mentioned in the introduction, when performing TST-based ligand screening, it is desirable to obtain a reaction path network and consider all important transition states that could be involved in the reaction. Therefore, we



(a) Hydroformylation of terminal olefin



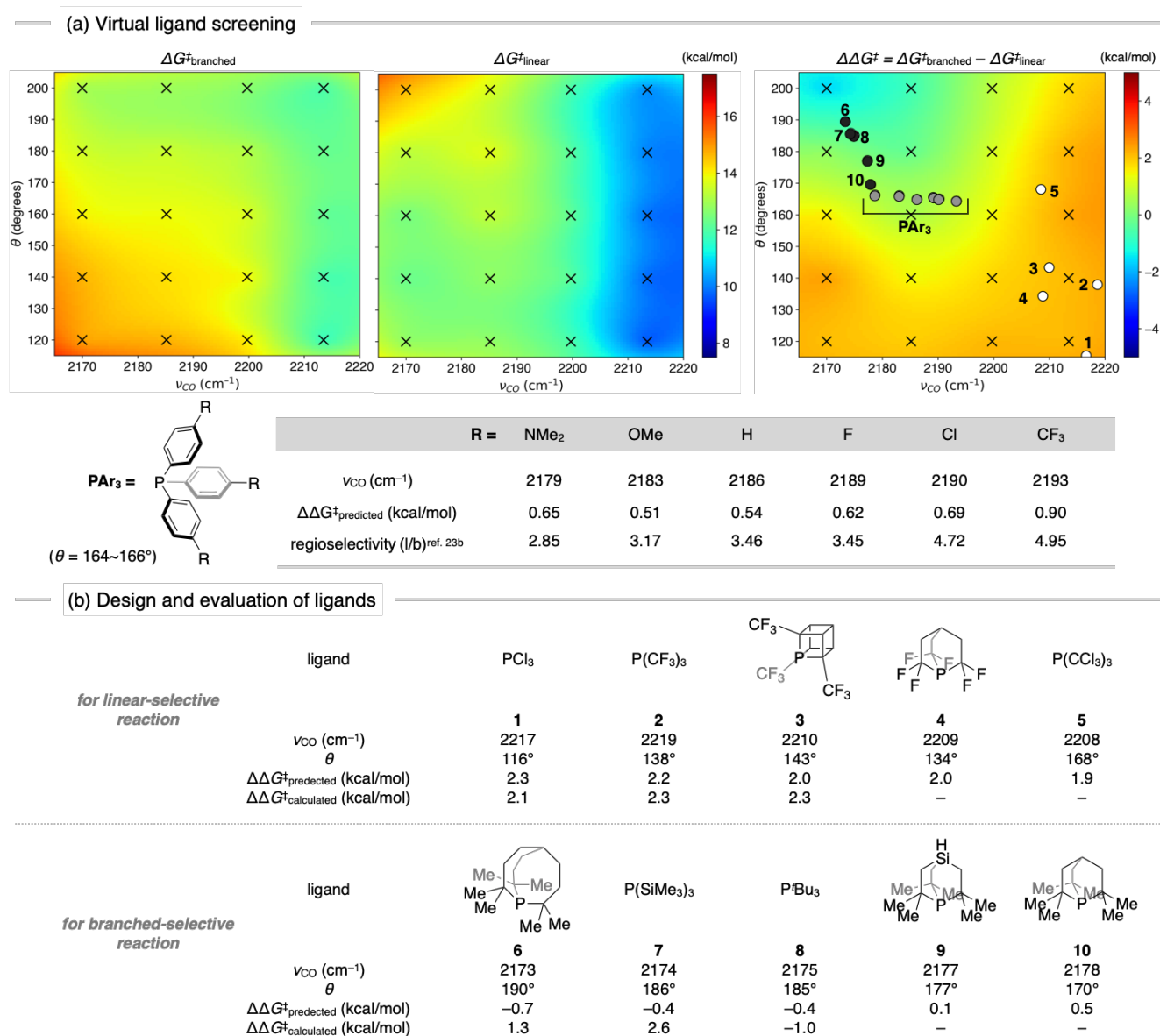
(b) Automated reaction path search



**Figure 8.** Hydroformylation of a terminal olefin and automated reaction path search for a model system to find the transition state that determines the regioselectivity.

performed an automated reaction path search using the single-component artificial force-induced reaction (SC-AFIR) method (Figure 8b).<sup>15</sup> The search was executed with a model system, in which 1-pentene and rhodium complex **G** with one virtual  $\text{PCl}^*_3$  ligand were employed as inputs at the B3LYP/LanL2DZ level (see Figures S9 and S10 for discussion of the case in which complex **G** has two phosphorus(III) ligands). In this calculation, the electronic and steric parameters of the virtual ligand were set to reproduce those of a representative phosphine ligand,  $\text{PPh}_3$ . The reaction path network shown in Figure 8b (left) was obtained, in which each node represents one locally stable structure (EQ) on the PES at the computational level adopted in the automated search and each edge corresponds to an approximated transition state (PT) connecting two EQs. Notably, the virtual ligand enables us to perform automated reaction path searches for a transition metal complex by avoiding the huge quantum chemical calculation cost associated with phosphorus(III) ligands (e.g., 34 atoms for

PPH<sub>3</sub>). The virtual ligand also reduces the number of conformational isomers for the phosphorus(III) ligand, which significantly reduces the gradient calculations required in the reaction path search. With the reaction path network in hand, we extracted the reaction paths corresponding to the insertion of 1-pentene and refined them at the  $\omega$ B97X-D/SDD, 6-31+G\* level (Figure 8b, right). Insertion was found to proceed from trigonal bipyramidal complex (*e*)-**H** or (*a*)-**H** to obtain four alkylrhodium complexes (*trans*-**I**, *cis*-**I**, *trans*-**J**, and *cis*-**J**). From (*e*)-**H**, in which the virtual ligand occupies an equatorial position, four different reaction paths resulting in each of the alkylrhodium complexes were found along with the corresponding transition states (**TS<sub>HI</sub>-1**, **TS<sub>HI</sub>-2**, **TS<sub>HJ</sub>-1**, and **TS<sub>HJ</sub>-2**, see Figure S11 for their geometries). On the other hand, from (*a*)-**H**, in which the virtual ligand is located at an apical position, only two reaction paths were obtained, which were connected to *cis*-**I** and *cis*-**J** through **TS<sub>HI</sub>-3** and **TS<sub>HJ</sub>-3**, respectively (see Figure S11 for their geometries). Based on these results,



**Figure 9.** Design of phosphorus(III) ligands for regioselective hydroformylation based on VLA screening and computational validation.

we supposed that the activation energies for the insertion of 1-pentene to give branched alkylrhodium **I** ( $\Delta G^\ddagger_{\text{branched}}$ ) and linear analogue **J** ( $\Delta G^\ddagger_{\text{linear}}$ ) can be given as follows:

$$\Delta G^\ddagger_{\text{branched}} = \min\{G(\text{TS}_{\text{HI}}-1), G(\text{TS}_{\text{HI}}-2), G(\text{TS}_{\text{HI}}-3)\} \\ - \min\{G((e)\text{-H}), G((a)\text{-H})\}$$

$$\Delta G^\ddagger_{\text{linear}} = \min\{G(\text{TS}_{\text{HJ}}-1), G(\text{TS}_{\text{HJ}}-2), G(\text{TS}_{\text{HJ}}-3)\} \\ - \min\{G((e)\text{-H}), G((a)\text{-H})\}$$

where  $G(\mathbf{X})$  is the Gibbs energy of **X**. With the current parameters of the virtual ligand ( $r_0 = 1.62 \text{ \AA}$  and  $\theta = 165^\circ$ ), for example,  $\Delta G^\ddagger_{\text{branched}}$  and  $\Delta G^\ddagger_{\text{linear}}$  were calculated to be 13.8 and 13.5 kcal/mol, respectively.

Considering these results, we examined the effect of the electronic and steric parameters of the virtual ligand on the  $\Delta G^\ddagger_{\text{branched}}$  and  $\Delta G^\ddagger_{\text{linear}}$  values as well as the regioselectivity of the reaction (Figure 9a). Based on the above-obtained structures, geometry optimization and harmonic vibrational frequency analyses of  $(e)\text{-H}$ ,  $(a)\text{-H}$ , **TS<sub>HI</sub>-1**, **TS<sub>HI</sub>-**

**2**, **TS<sub>HI</sub>-3**, **TS<sub>HJ</sub>-1**, **TS<sub>HJ</sub>-2**, and **TS<sub>HJ</sub>-3** were performed for 20 virtual systems ( $r_0 = 1.4, 1.6, 1.8$ , or  $2.0 \text{ \AA}$ , and  $\theta = 120, 140, 160, 180$ , or  $200^\circ$ ). In these calculations, several transition state geometries were not obtained. Hence, the  $\Delta G^\ddagger_{\text{branched}}$  and  $\Delta G^\ddagger_{\text{linear}}$  values were calculated according to the equations above assuming that these transition states are not energetically favored. The obtained results are plotted against the electronic and steric parameters of the virtual ligand, and the contour maps in Figure 9a were constructed. The left and middle contour maps represent the effects of the electronic and steric parameters on  $\Delta G^\ddagger_{\text{branched}}$  and  $\Delta G^\ddagger_{\text{linear}}$ , whereas the right contour map corresponds to the regioselectivity ( $\Delta\Delta G^\ddagger = \Delta G^\ddagger_{\text{branched}} - \Delta G^\ddagger_{\text{linear}}$ ) of the reaction. Although the energies predicted with the virtual ligand might only be qualitatively correct, the contour maps can indicate trends.  $\Delta\Delta G^\ddagger$  was found to be positive in almost all the examined parameter space, implying the preferential formation of the linear product. This result is consistent with reported experimental results, where the linear

product is the major product in the hydroformylation of unactivated terminal olefins such as 1-pentene.<sup>23</sup> Furthermore, for the hydroformylation of 1-hexene using PPh<sub>3</sub> and its derivatives (P(*p*-R-C<sub>6</sub>H<sub>4</sub>)<sub>3</sub>, R = NMe<sub>2</sub>, OMe, F, Cl, and CF<sub>3</sub>), Moser et al. have shown experimentally that the ratio of linear products increases as the electron-donating ability of the ligand decreases.<sup>23b</sup> The contour map for  $\Delta\Delta G^\ddagger$  also reproduced this experimental result. Based on the TEP values and cone angles of these triarylphosphines, the  $\Delta\Delta G^\ddagger$  value for each ligand was estimated using the contour map (gray circles, Figure 9a). The  $\Delta\Delta G^\ddagger$  value (or linear selectivity of the reaction) was found to increase with an increase in the TEP value (or decrease in the electron-donating ability), except for that of P(*p*-NMe<sub>2</sub>-C<sub>6</sub>H<sub>4</sub>)<sub>3</sub>, which was also an outlier in Moser's report (table, Figure 9a).<sup>23b</sup>

Based on the VLA screening results, we then attempted to design phosphorus(III) ligands that show high regioselectivity. According to the contour maps in Figure 9a, ligands with high  $v_{CO}$  and small  $\theta$  values should result in low activation barriers ( $\Delta G^\ddagger_{\text{branched}}$  and  $\Delta G^\ddagger_{\text{linear}}$ ) and high linear selectivity ( $\Delta\Delta G^\ddagger$ ). Therefore, we designed five phosphorus(III) ligand candidates that are expected to meet these criteria (**1–5**, Figure 9b). It should be noted that, taking advantage of the in silico screening, the stability and synthetic accessibility of the candidates were not taken into consideration. The  $\Delta\Delta G^\ddagger$  value for each candidate was then predicted using the contour map in Figure 9a (white circles) based on the TEP value and cone angle ( $\Delta\Delta G^\ddagger_{\text{predicted}}$ , Figure 9b). As expected, all these candidates were predicted to have positive  $\Delta\Delta G^\ddagger$  values, implying the preferential formation of the linear product. To validate these results computationally, ligands **1–3**, which were predicted to have the largest  $\Delta\Delta G^\ddagger$  values, were chosen, and geometry optimization of the corresponding structures to (*e*)-**H**, (*a*)-**H**, **TS<sub>H</sub>I-1**, **TS<sub>H</sub>I-2**, **TS<sub>H</sub>I-3**, **TS<sub>H</sub>J-1**, **TS<sub>H</sub>J-2** and **TS<sub>H</sub>J-3** was conducted with real ligands. The  $\Delta\Delta G^\ddagger$  values were then calculated based on these results ( $\Delta\Delta G^\ddagger_{\text{calculated}}$ , Figure 9b). Good agreement was confirmed between  $\Delta\Delta G^\ddagger_{\text{predicted}}$  and  $\Delta\Delta G^\ddagger_{\text{calculated}}$ , suggesting that ligands **1–3** have potential for the linear-selective hydroformylation of terminal olefins.

Similarly, we designed phosphorus(III) ligands that should enable branched-selective hydroformylation. As described above, the hydroformylation of unactivated terminal olefins generally affords linear isomers as the major product, with few reports on the branched-selective hydroformylation of these compounds.<sup>24</sup> Nevertheless, the contour map in Figure 9a implies that negative  $\Delta\Delta G^\ddagger$  values can be achieved when electron-rich and bulky ligands are employed, resulting in branched-selective reactions. With the aim of developing ligands for branched-selective reactions, we designed phosphorus(III) ligands **6–10**. The  $\Delta\Delta G^\ddagger$  values of these ligands were predicted based on their TEP values and cone angles (black circles, Figure 9a and  $\Delta\Delta G^\ddagger_{\text{predicted}}$ , Figure 9b), and these results were confirmed by real system calculations for ligands **6–8** ( $\Delta\Delta G^\ddagger_{\text{calculated}}$ , Figure 9b). The  $\Delta\Delta G^\ddagger_{\text{calculated}}$  value for **8** was negative (−1.0 kcal/mol), whereas those of **6** and **7** were positive. Consequently, ligand **8** was suggested to have potential for the branched-selective hydroformylation of unactivated terminal olefins such as 1-pentene.

## 4. CONCLUSIONS

In summary, we proposed a virtual-ligand-assisted (VLA) screening method to realize practical TST-based ligand screening for the development of transition metal catalysis and catalysts. To achieve this, we developed PCl<sub>3</sub> as a virtual ligand that parameterizes the electronic and steric effects of phosphorus(III) ligands. This virtual ligand not only reduced the computational resources required for quantum chemical calculations of transition metal complexes but also enabled a parameter-based exploration of phosphorus(III) ligands by optimizing the electronic and steric parameters of the reaction of interest. In combination with an automated reaction path search using the SC-AFIR method, we demonstrated VLA screening for the regioselective hydroformylation of terminal olefins, successfully designing phosphorus(III) ligands that can potentially achieve linear- or branched-selective hydroformylation. As VLA screening can be conducted in silico without any experimental work, this approach may be able to streamline ligand screening, which is the most important but time- and resource-consuming process in the development of transition metal catalysis. Attempts to improve the virtual ligand model to enable quantitative predictions and to apply this concept to other representative classes of organic ligands, such as bidentate phosphines and NHCs, are currently in progress.

## ASSOCIATED CONTENT

**Supporting Information.** Computational details, list of calculated phosphorus(III) ligands, Tolman's maps (corresponding to Figures 6b, 6c, 7b, and 7c) with all ligands assigned, discussion of terminal olefin hydroformylation, and cartesian coordinates of optimized structures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\* [matsuoka.wataru@sci.hokudai.ac.jp](mailto:matsuoka.wataru@sci.hokudai.ac.jp) (W.M.)

\* [smaeda@eis.hokudai.ac.jp](mailto:smaeda@eis.hokudai.ac.jp) (S.M.)

### Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENT

This work was supported by JST via ERATO grant JPMJER1903. Support was also provided by the Institute for Chemical Reaction Design and Discovery (ICReDD), which was established by the World Premier International Research Initiative (WPI), MEXT, Japan.

## REFERENCES

- (1) For selected reviews, see: (a) Beller, M.; Cornils, B.; Frohning, C. D.; Kohlpaintner, C. W. Progress in Hydroformylation and Carbonylation. *J. Mol. Catal. A: Chem.* **1995**, *104*, 17–85. (b) Agbossou, F.; Carpentier, J.-F.; Mortreux, A. Asymmetric Hydroformylation. *Chem. Rev.* **1995**, *95*, 2485–2506. (c) Breit, B. Synthetic Aspects of Stereoselective Hy-

- droformylation. *Acc. Chem. Res.* **2003**, *36*, 264–275. (d) Franke, R.; Selent, D.; Börner, A. Applied Hydroformylation. *Chem. Rev.* **2012**, *112*, 5675–5732. (e) Chakraborty, S.; Almasalma, A. A.; de Vries, J. G. Recent Developments in Asymmetric Hydroformylation. *Catal. Sci. Technol.* **2021**, *11*, 5388–5411.
- (2) For selected reviews, see: (a) Schrock, R. R.; Hoveyda, A. H. Molybdenum and Tungsten Imido Alkylidene Complexes as Efficient Olefin-Metathesis Catalysts. *Angew. Chem., Int. Ed.* **2003**, *42*, 4592–4633. (b) Grubbs, R. H. Olefin Metathesis. *Tetrahedron* **2004**, *60*, 7117–7140. (c) Donohoe, T. J.; Orr, A. J.; Bingham, M. Ring-Closing Metathesis as a Basis for the Construction of Aromatic Compounds. *Angew. Chem., Int. Ed.* **2006**, *45*, 2664–2670. (d) Hoveyda, A. H.; Zhugralin, A. R. The Remarkable Metal-Catalysed Olefin Metathesis Reaction. *Nature* **2007**, *450*, 243–251. (e) Nolan, S. P.; Clavier, H. Chemoselective Olefin Metathesis Transformations Mediated by Ruthenium Complexes. *Chem. Soc. Rev.* **2010**, *39*, 3305–3316. (f) Vougioukalakis, G. C.; Grubbs, R. H. Ruthenium-Based Heterocyclic Carbene-Coordinated Olefin Metathesis Catalysts. *Chem. Rev.* **2010**, *110*, 1746–1787. (g) Ogbay, O. M.; Warner, N. C.; O’Leary, D. J.; Grubbs, R. H. Recent Advances in Ruthenium-Based Olefin Metathesis. *Chem. Soc. Rev.* **2018**, *47*, 4510–4544.
- (3) For selected reviews, see: (a) Noyori, R.; Hashiguchi, S. Asymmetric Transfer Hydrogenation Catalyzed by Chiral Ruthenium Complexes. *Acc. Chem. Res.* **1997**, *30*, 97–102. (b) Noyori, R.; Ohkuma, T. Asymmetric Catalysis by Architectural and Functional Molecular Engineering: Practical Chemo- and Stereoselective Hydrogenation of Ketones. *Angew. Chem., Int. Ed.* **2001**, *40*, 40–73. (c) Xie, J.-H.; Zhu, S.-F.; Zhou, Q.-L. Transition Metal-Catalyzed Enantioselective Hydrogenation of Enamines and Imines. *Chem. Rev.* **2011**, *111*, 1713–1760. (d) Wang, D.-S.; Chen, Q.-A.; Lu, S.-M.; Zhou, Y.-G. Asymmetric Hydrogenation of Heteroarenes and Arenes. *Chem. Rev.* **2012**, *112*, 2557–2590. (e) Wang, D.; Astruc, D. The Golden Age of Transfer Hydrogenation. *Chem. Rev.* **2015**, *115*, 6621–6686. (f) Wen, J.; Wang, F.; Zhang, X. Asymmetric Hydrogenation Catalyzed by First-Row Transition Metal Complexes. *Chem. Soc. Rev.* **2021**, *50*, 3211–3237.
- (4) For selected reviews, see: (a) Miyauchi, N.; Suzuki, A. Palladium-Catalyzed Cross-Coupling Reactions of Organoboron Compounds. *Chem. Rev.* **1995**, *95*, 2457–2483. (b) Johansson Seechurn, C. C. C.; Kitching, M. O.; Colacot, T. J.; Snieckus, V. Palladium-Catalyzed Cross-Coupling: A Historical Contextual Perspective to the 2010 Nobel Prize. *Angew. Chem., Int. Ed.* **2012**, *51*, 5062–5085. (c) Magano, J.; Dunetz, J. R. Large-Scale Applications of Transition Metal-Catalyzed Couplings for the Synthesis of Pharmaceuticals. *Chem. Rev.* **2011**, *111*, 2177–2250. (d) Yeung, C. S.; Dong, V. M. Catalytic Dehydrogenative Cross-Coupling: Forming Carbon–Carbon Bonds by Oxidizing Two Carbon–Hydrogen Bonds. *Chem. Rev.* **2011**, *111*, 1215–1292. (e) Ruiz-Castillo, P.; Buchwald, S. L. Applications of Palladium-Catalyzed C–N Cross-Coupling Reactions. *Chem. Rev.* **2016**, *116*, 12564–12649. (f) Zhang, Y.-F.; Shi, Z.-J. Upgrading Cross-Coupling Reactions for Biaryl Syntheses. *Acc. Chem. Res.* **2019**, *52*, 161–169.
- (5) For selected reviews focusing on specific classes of organic ligands, see: (a) Berthod, M.; Mignani, G.; Woodward, G.; Lemaire, M. Modified BINAP: The How and the Why. *Chem. Rev.* **2005**, *105*, 1801–1836. (b) Arrayás, R. G.; Adrio, J.; Carretero, J. C. Recent Applications of Chiral Ferrocene Ligands in Asymmetric Catalysis. *Angew. Chem., Int. Ed.* **2006**, *45*, 7674–7715. (c) Martin, R.; Buchwald, S. L. Palladium-Catalyzed Suzuki–Miyaura Cross-Coupling Reactions Employing Dialkylbiaryl Phosphine Ligands. *Acc. Chem. Res.* **2008**, *41*, 1461–1473. (d) Valente, C.; Çalimsiz, S.; Hoi, K. H.; Mallik, D.; Sayah, M.; Organ, M. G. The Development of Bulky Palladium NHC Complexes for the Most-Challenging Cross-Coupling Reactions. *Angew. Chem., Int. Ed.* **2012**, *51*, 3314–3332. (e) Ye, B.; Cramer, N. Chiral Cyclopentadienyls: Enabling Ligands for Asymmetric Rh(III)-Catalyzed C–H Functionalizations. *Acc. Chem. Res.* **2015**, *48*, 1308–1318. (f) Zhao, Q.; Meng, G.; Nolan, S. P.; Szostak, M. N-Heterocyclic Carbene Complexes in C–H Activation Reactions. *Chem. Rev.* **2020**, *120*, 1981–2048.
- (6) For selected reviews and databases focusing on the steric and electronic effects of organic ligands, see: (a) Tolman, C. A. Steric Effects of Phosphorus Ligands in Organometallic Chemistry and Homogeneous Catalysis. *Chem. Rev.* **1977**, *77*, 313–348. (b) Fey, N.; Tsipis, A. C.; Harris, S. E.; Harvey, J. N.; Orpen, A. G.; Mansson, R. A. Development of a Ligand Knowledge Base, Part 1: Computational Descriptors for Phosphorus Donor Ligands. *Chem. Eur. J.* **2006**, *12*, 291–302. (c) Díez-González, S.; Nolan, S. P. Stereoelectronic Parameters Associated with N-Heterocyclic Carbene (NHC) Ligands: A Quest for Understanding. *Coord. Chem. Rev.* **2007**, *251*, 874–883. (d) Fey, N.; Orpen, A. G.; Harvey, J. N. Building Ligand Knowledge Bases for Organometallic Chemistry: Computational Description of Phosphorus(III)-Donor Ligands and the Metal–Phosphorus Bond. *Coord. Chem. Rev.* **2009**, *253*, 704–722. (e) Jover, J.; Fey, N.; Harvey, J. N.; Lloyd-Jones, G. C.; Orpen, A. G.; Owen-Smith, G. J. J.; Murray, P.; Hose, D. R. J.; Osborne, R.; Purdie, M. Expansion of the Ligand Knowledge Base for Monodentate P-Donor Ligands (LKB-P). *Organometallics* **2010**, *29*, 6245–6258. (f) Jover, J.; Fey, N.; Harvey, J. N.; Lloyd-Jones, G. C.; Orpen, A. G.; Owen-Smith, G. J. J.; Murray, P.; Hose, D. R. J.; Osborne, R.; Purdie, M. Expansion of the Ligand Knowledge Base for Chelating P,P-Donor Ligands (LKB-PP). *Organometallics* **2012**, *31*, 5302–5306. (g) Durand, D. J.; Fey, N. Computational Ligand Descriptors for Catalyst Design. *Chem. Rev.* **2019**, *119*, 6561–6594.
- (7) For selected examples of electronic descriptors, see: (a) Tolman, C. A. Electron Donor–Acceptor Properties of Phosphorus Ligands. Substituent Additivity. *J. Am. Chem. Soc.* **1970**, *92*, 2953–2956. (b) Perrin, L.; Clot, E.; Eisenstein, O.; Loch, J.; Crabtree, R. H. Computed Ligand Electronic Parameters from Quantum Chemistry and Their Relation to Tolman Parameters, Lever Parameters, and Hammett Constants. *Inorg. Chem.* **2001**, *40*, 5806–5811. (c) Gusev, D. G. Donor Properties of a Series of Two-Electron Ligands. *Organometallics* **2009**, *28*, 763–770. (d) Kalescky, R.; Kraka, E.; Cremer, D. New Approach to Tolman’s Electronic Parameter Based on Local Vibrational Modes. *Inorg. Chem.* **2014**, *53*, 478–495. (e) Suresh, C. H.; Koga, N. Quantifying the Electronic Effect of Substituted Phosphine Ligands via Molecular Electrostatic Potential. *Inorg. Chem.* **2002**, *41*, 1573–1578. (f) Coll, D. S.; Vidal, A. B.; Rodríguez, J. A.; Ocampo-Mavárez, E.; Añez, R.; Sierraalta, A. A Simple Method for the Determination of the Tolman Electronic Parameter of Different Phosphorus Containing Ligands, by Means of the Average Local Ionization Energy. *Inorg. Chim. Acta* **2015**, *436*, 163–168. (g) Setiawan, D.; Kalescky, R.; Kraka, E.; Cremer, D. Direct Measure of Metal–Ligand Bonding Replacing the Tolman Electronic Parameter. *Inorg. Chem.* **2016**, *55*, 2332–2344.
- (8) For selected examples of steric descriptors, see: (a) Guzei, I. A.; Wendt, M. An Improved Method for the Computation of Ligand Steric Effects Based on Solid Angles. *Dalton Trans.* **2006**, 3991–3999. (b) Clavier, H.; Nolan, S. P. Percent Buried Volume for Phosphine and N-Heterocyclic Carbene Ligands: Steric Properties in Organometallic Chemistry. *Chem. Commun.* **2010**, *46*, 841–861. (c) Bilibrey, J. A.; Kazez, A. H.; Locklin, J.; Allen, W. D. Exact Ligand Cone Angles. *J. Comput. Chem.* **2013**, *34*, 1189–1197. (d) van Leeuwen, P. W. N. M.; Kamer, P. C. J.; Reek, J. N. H.; Dierkes, P. Ligand Bite Angle Effects in Metal-Catalyzed C–C Bond Formation. *Chem. Rev.* **2000**, *100*, 2741–2769. See also ref. 6a, b, and f.



- (9) For selected reviews focusing on chemical libraries or the high-throughput screening of ligands, see: (a) Goudriaan, P. E.; van Leeuwen, P. W. N. M.; Birkholz, M.-N.; Reek, J. N. H. Libraries of Bidentate Phosphorus Ligands; Synthesis Strategies and Application in Catalysis. *Eur. J. Inorg. Chem.* **2008**, 2008, 2939–2958. (b) Collins, K. D.; Gensch, T.; Glorius, F. Contemporary Screening Approaches to Reaction Discovery and Development. *Nat. Chem.* **2014**, 6, 859–871. (c) Renom-Carrasco, M.; Lefort, L. Ligand Libraries for High Throughput Screening of Homogeneous Catalysts. *Chem. Soc. Rev.* **2018**, 47, 5038–5060.
- (10) For selected examples of the ligand screening by informatics approach, see: (a) Niemeyer, Z. L.; Milo, A.; Hickey, D. P.; Sigman, M. S. Parameterization of Phosphine Ligands Reveals Mechanistic Pathways and Predicts Reaction Outcomes. *Nat. Chem.* **2016**, 8, 610–617. (b) Wu, K.; Doyle, A. G. Parameterization of Phosphine Ligands Demonstrates Enhancement of Nickel Catalysis via Remote Steric Effects. *Nat. Chem.* **2017**, 9, 779–784. (c) Zahrt, A. F.; Henle, J. J.; Rose, B. T.; Wang, Y.; Darrow, W. T.; Denmark, S. E. Prediction of Higher-Selectivity Catalysts by Computer-Driven Workflow and Machine Learning. *Science* **2019**, 363, eaau5631. (d) Brethomé, A. V.; Paton, R. S.; Fletcher, S. P. Retooling Asymmetric Conjugate Additions for Sterically Demanding Substrates with an Iterative Data-Driven Approach. *ACS Catal.* **2019**, 9, 7179–7187. (e) Hueffel, J. A.; Sperger, T.; Funes-Ardoiz, I.; Ward, J. S.; Rissanen, K.; Schoenebeck, F. Accelerated Dinuclear Palladium Catalyst Identification through Unsupervised Machine Learning. *Science* **2021**, 374, 1134–1140. (f) Gensch, T.; dos Passos Gomes, G.; Friederich, P.; Peters, E.; Gaudin, T.; Pollice, R.; Jorner, K.; Nigam, A.; Lindner-D'Addario, M.; Sigman, M. S.; Aspuru-Guzik, A. A Comprehensive Discovery Platform for Organophosphorus Ligands for Catalysis. *ChemRxiv* **2021**, 10.26434/chemrxiv.12996665.v1.
- (11) (a) Houk, K. N.; Liu, F. Holy Grails for Computational Organic Chemistry and Biochemistry. *Acc. Chem. Res.* **2017**, 50, 539–543. (b) Hammes-Schiffer, S. Catalysts by Design: The Power of Theory. *Acc. Chem. Res.* **2017**, 50, 561–566. (c) Poree, C.; Schoenebeck, F. A Holy Grail in Chemistry: Computational Catalyst Design: Feasible or Fiction? *Acc. Chem. Res.* **2017**, 50, 605–608.
- (12) For a review, see: (a) Ahn, S.; Hong, M.; Sundararajan, M.; Ess, D. H.; Baik, M.-H. Design and Optimization of Catalysts Based on Mechanistic Insights Derived from Quantum Chemical Reaction Modeling. *Chem. Rev.* **2019**, 119, 6509–6560. For selected examples, see: (b) Nielsen, M. C.; Bonney, K. J.; Schoenebeck, F. Computational Ligand Design for the Reductive Elimination of  $\text{ArCF}_3$  from a Small Bite Angle  $\text{Pd}^{\text{II}}$  Complex: Remarkable Effect of a Perfluoroalkyl Phosphine. *Angew. Chem., Int. Ed.* **2014**, 53, 5903–5906. (c) Straker, R. N.; Peng, Q.; Mekareeya, A.; Paton, R. S.; Anderson, E. A. Computational Ligand Design in Enantio- and Diastereoselective Ynamide [5 + 2] Cycloisomerization. *Nat. Commun.* **2016**, 7, 10109. (d) Korenaga, T.; Sasaki, R.; Takemoto, T.; Yasuda, T.; Watanabe, M. Computationally-Led Ligand Modification Using Interplay between Theory and Experiments: Highly Active Chiral Rhodium Catalyst Controlled by Electronic Effects and  $\text{CH}-\pi$  Interactions. *Adv. Synth. Catal.* **2018**, 360, 322–333.
- (13) For selected reviews focusing on automated or high-throughput computations, see: (a) Foscatto, M.; Jensen, V. R. Automated in Silico Design of Homogeneous Catalysts. *ACS Catal.* **2020**, 10, 2354–2377. (b) Nandy, A.; Duan, C.; Taylor, M. G.; Liu, F.; Steeves, A. H.; Kulik, H. J. Computational Discovery of Transition-Metal Complexes: From High-Throughput Screening to Machine Learning. *Chem. Rev.* **2021**, 121, 9927–10000. For selected examples, see: (c) Ioannidis, E. I.; Gani, T. Z. H.; Kulik, H. J. molSimplify: A Toolkit for Automating Discovery in Inorganic Chemistry. *J. Comput. Chem.* **2016**, 37, 2106–2117. (d) Guan, Y.; Ingman, V. M.; Rooks, B. J.; Wheeler, S. E. AARON: An Automated Reaction Optimizer for New Catalysts. *J. Chem. Theory Comput.* **2018**, 14, 5249–5261. (e) Rosales, A. R.; Wahlers, J.; Limé, E.; Meadows, R. E.; Leslie, K. W.; Savin, R.; Bell, F.; Hansen, E.; Helquist, P.; Munday, R. H.; Wiest, O.; Norrby, P.-O. Rapid Virtual Screening of Enantioselective Catalysts Using CatVS. *Nat. Catal.* **2019**, 2, 41–45.
- (14) (a) Friederich, P.; dos Passos Gomes, G.; De Bin, R.; Aspuru-Guzik, A.; Balcells, D. Machine Learning Dihydrogen Activation in the Chemical Space Surrounding Vaska's Complex. *Chem. Sci.* **2020**, 11, 4584–4601. (b) Maley, S. M.; Kwon, D.-H.; Rollins, N.; Stanley, J. C.; Sydora, O. L.; Bischof, S. M.; Ess, D. H. Quantum-Mechanical Transition-State Model Combined with Machine Learning Provides Catalyst Design Features for Selective Cr Olefin Oligomerization. *Chem. Sci.* **2020**, 11, 9665–9674. See also ref. 10f.
- (15) (a) Maeda, S.; Taketsugu, T.; Morokuma, K. Exploring Transition State Structures for Intramolecular Pathways by the Artificial Force Induced Reaction Method. *J. Comput. Chem.* **2014**, 35, 166–173. (b) Maeda, S.; Harabuchi, Y.; Takagi, M.; Saita, K.; Suzuki, K.; Ichino, T.; Sumiya, Y.; Sugiyama, K.; Ono, Y. Implementation and Performance of the Artificial Force Induced Reaction Method in the GRRM17 Program. *J. Comput. Chem.* **2018**, 39, 233–250. (c) Maeda, S.; Harabuchi, Y. Exploring Paths of Chemical Transformations in Molecular and Periodic Systems: An Approach Utilizing Force. *WIREs Comput. Mol. Sci.* **2021**, 11, No. e1538.
- (16) (a) Behn, A.; Zimmerman, P. M.; Bell, A. T.; Head-Gordon, M. Efficient Exploration of Reaction Paths via a Freezing String Method. *J. Chem. Phys.* **2011**, 135, 224108. (b) Suleimanov, Y. V.; Green, W. H. Automated Discovery of Elementary Chemical Reaction Steps Using Freezing String and Berny Optimization Methods. *J. Chem. Theory Comput.* **2015**, 11, 4248–4259. (c) Mallikarjun Sharada, S.; Zimmerman, P. M.; Bell, A. T.; Head-Gordon, M. Automated Transition State Searches without Evaluating the Hessian. *J. Chem. Theory Comput.* **2012**, 8, 5166–5174. (d) Zimmerman, P. M. Automated Discovery of Chemically Reasonable Elementary Reaction Steps. *J. Comput. Chem.* **2013**, 34, 1385–1392. (e) Zimmerman, P. M. Single-Ended Transition State Finding with the Growing String Method. *J. Comput. Chem.* **2015**, 36, 601–611. (f) Zádor, J.; Najm, H. N. In *KinBot 1.0: a Code for Automatic PES Exploration*; Report No. SAND2012-8095; Sandia National Laboratories: 2013.
- (17) Ohnishi, Y.; Nakao, Y.; Sato, H.; Sakaki, S. Frontier Orbital Consistent Quantum Capping Potential (FOC-QCP) for Bulky Ligand of Transition Metal Complexes. *J. Phys. Chem. A* **2008**, 112, 1946–1955.
- (18) (a) Littke, A. F.; Fu, G. C. A Convenient and General Method for Pd-Catalyzed Suzuki Cross-Couplings of Aryl Chlorides and Arylboronic Acids. *Angew. Chem., Int. Ed.* **1998**, 37, 3387–3388. (b) Ahlquist, M.; Norrby, P.-O. Oxidative Addition of Aryl Chlorides to Monoligated Palladium(0): A DFT-SCRF Study. *Organometallics* **2007**, 26, 550–553. (c) Schoenebeck, F.; Houk, K. N. Ligand-Controlled Regioselectivity in Palladium-Catalyzed Cross Coupling Reactions. *J. Am. Chem. Soc.* **2010**, 132, 2496–2497. (d) McMullin, C. L.; Fey, N.; Harvey, J. N. Computed Ligand Effects on the Oxidative Addition of Phenyl Halides to Phosphine Supported Palladium(0) Catalysts. *Dalton Trans.* **2014**, 43, 13545–13556.
- (19) (a) Campeau, L.-C.; Parisien, M.; Leblanc, M.; Fagnou, K. Biaryl Synthesis via Direct Arylation: Establishment of an Efficient Catalyst for Intramolecular Processes. *J. Am. Chem. Soc.* **2004**, 126, 9186–9187. (b) Lafrance, M.; Lapointe, D.; Fagnou, K. Mild and Efficient Palladium-Catalyzed Intramolecular Direct Arylation Reactions. *Tetrahedron* **2008**, 64, 6015–6020. (c) Korenaga, T.; Suzuki, N.; Sueda, M.; Shimada,

- K. Ligand Effect on Direct Arylation by CMD Process. *J. Organomet. Chem.* **2015**, 780, 63–69. (d) Davies, D. L.; Macgregor, S. A.; McMullin, C. L. Computational Studies of Carboxylate-Assisted C–H Activation and Functionalization at Group 8–10 Transition Metal Centers. *Chem. Rev.* **2017**, 117, 8649–8709.
- (20) Falivene, L.; Cao, Z.; Petta, A.; Serra, L.; Poater, A.; Oliva, R.; Scarano, V.; Cavallo, L. Towards the Online Computer-Aided Design of Catalytic Pockets. *Nat. Chem.* **2019**, 11, 872–879.
- (21) For selected reviews, see: (a) Neel, A. J.; Hilton, M. J.; Sigman, M. S.; Toste, F. D. Exploiting Non-Covalent  $\pi$  Interactions for Catalyst Design. *Nature* **2017**, 543, 637–646. (b) Davis, H. J.; Phipps, R. J. Harnessing Non-Covalent Interactions to Exert Control over Regioselectivity and Site-Selectivity in Catalytic Reactions. *Chem. Sci.* **2017**, 8, 864–877. (c) Proctor, R. S. J.; Colgan, A. C.; Phipps, R. J. Exploiting Attractive Non-Covalent Interactions for the Enantioselective Catalysis of Reactions Involving Radical Intermediates. *Nat. Chem.* **2020**, 12, 990–1004.
- (22) (a) Evans, D.; Osborn, J. A.; Wilkinson, G. Hydroformylation of Alkenes by Use of Rhodium Complex Catalysts. *J. Chem. Soc. A* **1968**, 3133–3142. (b) Decker, S. A.; Cundari, T. R. DFT Study of the Ethylene Hydroformylation Catalytic Cycle Employing a  $\text{HRh}(\text{PH}_3)_2(\text{CO})$  Model Catalyst. *Organometallics* **2001**, 20, 2827–2841. (c) Kumar, M.; Chaudhari, R. V.; Subramaniam, B.; Jackson, T. A. Ligand Effects on the Regioselectivity of Rhodium-Catalyzed Hydroformylation: Density Functional Calculations Illuminate the Role of Long-Range Noncovalent Interactions. *Organometallics* **2014**, 33, 4183–4191.
- (23) (a) Brown, C. K.; Wilkinson, G. Homogeneous Hydroformylation of Alkenes with Hydridocarbonyltris(triphenylphosphine)rhodium(I) as Catalyst. *J. Chem. Soc. A* **1970**, 2753–2764. (b) Moser, W. R.; Papile, C. J.; Brannon, D. A.; Duwell, R. A.; Weininger, S. J. The Mechanism of Phosphine-Modified Rhodium-Catalyzed Hydroformylation Studied by CIR-FTIR. *J. Mol. Catal.* **1987**, 41, 271–292.
- (24) For selected examples, see: (a) Slagt, V. F.; Kamer, P. C. J.; van Leeuwen, P. W. N. M.; Reek, J. N. H. Encapsulation of Transition Metal Catalysts by Ligand-Template Directed Assembly. *J. Am. Chem. Soc.* **2004**, 126, 1526–1536. (b) Dabawala, A. A.; Jasra, R. V.; Bajaj, H. C. Selective Hydroformylation of 1-Hexene to Branched Aldehydes Using Rhodium Complex of Modified Bulky Phosphine and Phosphite Ligands. *Catal. Commun.* **2011**, 12, 403–407. (c) Noonan, G. M.; Fuentes, J. A.; Cobley, C. J.; Clarke, M. L. An Asymmetric Hydroformylation Catalyst that Delivers Branched Aldehydes from Alkyl Alkenes. *Angew. Chem., Int. Ed.* **2012**, 51, 2477–2480. (d) Phanopoulos, A.; Nozaki, K. Branched-Selective Hydroformylation of Nonactivated Olefins Using an N-Triphos/Rh Catalyst. *ACS Catal.* **2018**, 8, 5799–5809.

## For Table of Contents Only

