Virtual Ligand–Assisted Optimization: A Rational Strategy for Ligand Engineering

Wataru Matsuoka,*[a,b,c] Taihei Oki,[b,d] Ren Yamada,[e] Tomohiko Yokoyama,[d] Shinichi Suda,[e] Yu Harabuchi,[a,b] Satoru Iwata,*[a,b,d] Satoshi Maeda*[a,b,c]

[a]Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Kita 21, Nishi 10, Kita-ku, Sapporo, Hokkaido 001-0021, Japan. [b]JST, ERATO Maeda Artificial Intelligence in Chemical Reaction Design and Discovery Project, Kita 10, Nishi 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan. [c]Department of Chemistry, Faculty of Science, Hokkaido University, Kita 10, Nishi 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan. [d]Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8656, Japan. [e]Department of Chemical Science and Engineering, Graduate School of Science, Hokkaido University, Kita 10, Nishi 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan.

*Corresponding author:

E-mail: matsuoka.wataru@sci.hokudai.ac.jp, iwata@mist.i.u-tokyo.ac.jp, smaeda@eis.hokudai.ac.jp

ABSTRACT: Ligand engineering is one of the most important, but labor-intensive processes in the development of transition metal catalysis. Historically, this process has been streamlined by the invention of ligand descriptors such as Tolman's electronic parameter and the cone angle. Analyzing reaction outcomes in terms of these parameters has enabled chemists to find important factors for designing optimal ligands. However, typical strategies for these analyses largely rely on regression approaches, which often requires many experimental data to understand non-intuitive trends. Here, we introduce the virtual ligand-assisted optimization (VLAO) method, a novel computational approach for ligand engineering. In this method, important features of ligands are identified by simple mathematical operations on equilibrium structures and/or transition states of interest, and derivative values of arbitrary objective functions with respect to ligand parameters are obtained. These derivative values are then used as a guiding principle to optimize ligands within the parameter space. The VLAO method was demonstrated in the optimization of monodentate and bidentate phosphine ligands including asymmetric quinoxaline-based ligands. In addition, we successfully found a highly selective ligand for the α-selective hydrogermylation of a terminal ynamide according to the suggested design principle by the VLAO method. These results would imply the potential utility of the VLAO method in optimizing wide variety of ligands in transition metal catalysis.

INTRODUCTION

Ligand plays one of the most important roles in the transition metal catalysis. During the last few decades, many sophisticated ligands which enable unprecedented transformations with high efficiency have been developed.[1] Except for serendipity, development of such ligands depends on rational ligand engineering.[2] Ligand engineering is a process where a molecular structure of a ligand is optimized to maximize reaction yield, selectivity and/or rate. To streamline this process, various electronic and steric parameters which describe inherent properties of ligands have been introduced, and correlation between these parameters and outcomes of the reaction have been investigated (Figure 1a, left).[3-6] For historical example, Tolman and coworkers have conducted a pioneering work in this field, where Tolman's electronic parameter (v_{CO}) and the cone angle (θ) were introduced as metrics for electronic and steric properties of monodentate phosphorous(III) ligands.[3] They revealed that, by plotting a measurable function F (*i.e.*, reaction yield) against these parameters (so-called Tolman's plot), the electronic and steric characters of the reaction, which are

used as the guidelines for designing optimal ligands, can be understood as the slope of the function with respect to each parameter ($\Delta F/\Delta v_{CO}$ and $\Delta F/\Delta \theta$). Their research has established the fundamental strategy of ligand engineering, and most of recent ligand discoveries have been achieved by (at least unconsciously) analyzing reaction outcomes in terms of ligand parameters (*i.e.*, $\Delta F/\Delta p$, where p is an arbitrary parameter).[2] Notably, many advanced electronic/steric parameters have been recently developed, largely owing to the emergence of computational tools, enabling highly sophisticated design of ligand molecules.[4–6]

However, regardless of the simplicity of its fundamental strategy, ligand engineering is still a time-consuming and labor-intensive process. This stems from the difficulty in evaluating the effect of a single ligand parameter on a reaction outcome (Figure 1a, right). Because all electronic and steric parameters are inherently determined by the molecular structure of a ligand, modification of the molecular structure would perturb all parameters, rather than a single parameter of interest. Hence, even if structures of ligands are systematically modified (*e.g.*, PEt$_3$, P^{i}Pr$_3$ and P^{t}Bu$_3$), what is

observed from experimental results is not the effect of a single parameter of interest but the sum of contributions of all parameter changes (ΔF). In other words, it is not possible to directly evaluate the slope of the objective function because none of parameters (p) can be treated as an independent variable. Therefore, the slope needs to be guessed by comprehensively analyzing reaction outcomes obtained from evaluation of many ligand molecules.[3] It should be noted that correlation between F and p is not always simple, and hence a lot of time and resources are often required to collect a plenty amount of experimental data for understanding. Therefore, informatics techniques including regression analysis to rapidly capture important parameters for a given reaction are becoming promising tools in this field.[7,8]

Herein, we report the virtual ligand-assisted optimization (VLAO) method, a fundamentally new in silico approach for ligand engineering (Figure 1b). In this method, the electronic and steric characters of the reaction are estimated based on quantum chemical calculations and the following simple mathematical operations, rather than regression analysis of experimental data. This computational method is based on the use of the virtual ligand (VL),[9,10] which was introduced in our previous study about the virtual ligand-assisted screening (VLAS).[9] A VL is a dummy ligand for quantum chemical calculations to approximate electronic and steric properties of a real ligand. Whereas electronic and steric parameters of a real ligand are inherently determined by its chemical structure (see above), parameters in a VL (p_{VL}) can be independently tuned. In the VLAS

method, this feature of VLs was exploited to systematically investigate the entire parameter space to find an optimal parameter for the reaction of interest.[9] On the other hand, the VLAO method takes advantage of this feature to mathematically analyze electronic and steric characters of a reaction. As shown in Figure 1b (left), energies of an equilibrium structure (EQ) and transition state (TS) calculated using VLs can be differentiated in terms of parameters in the VLs ($dE_i/d p_{VL}$). Given that a reaction outcome such as reaction yield, selectivity or rate can be described as a function of energies of EQ and TS,[11,12] we can calculate derivative values of such an objective function in terms of VL parameters ($dF/d p_{VL}$) using $dE_i/d p_{VL}$. This value is exactly the extreme forms of $\Delta F/\Delta p$, which is pursued in typical ligand engineering study.[2,7,8] Hence, this would suggest a rational direction in which VL parameters should be tuned to maximize (minimize) the objective function (F). Then, we report a gradient-driven optimization of ligands within the VL parameter space (Figure 1b, right). Based on calculated derivative values ($dF/d p_{VL}$), the parameters in a VL are optimized by conjugate gradient (CG) method to maximize (minimize) the objective function F. This results in a VL with optimal parameters to achieve the target reaction most efficiently, which can be used as a rational guideline for designing optimal ligand molecules. The VLAO strategy was firstly demonstrated in three reported examples of ligand engineering, where monodentate, bidentate or asymmetric phosphine ligands are optimized. Then, the VLAO strategy was employed to experimentally find a highly selective ligand for the α-selective hydrogermylation of a terminal ynamide.
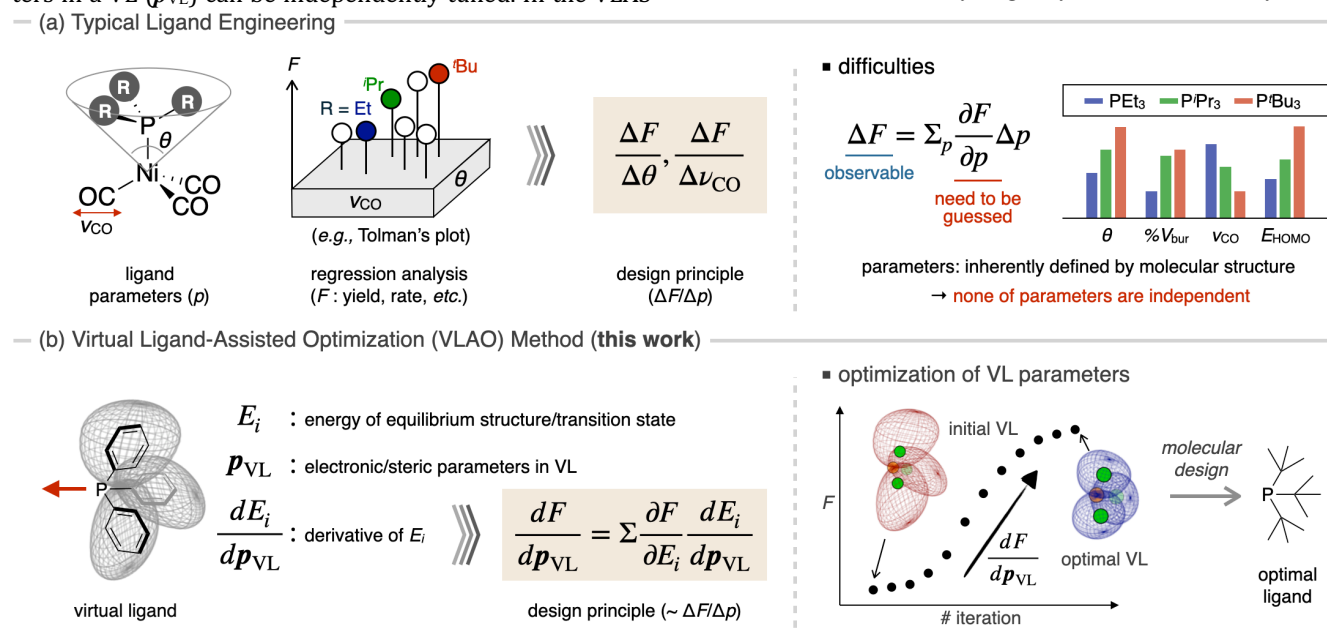

Figure 1. Schematic illustration of ligand engineering processes. (a) The fundamental strategy and difficulties in typical ligand engineering. (b) The strategy in the virtual ligand-assisted optimization method.

THEORY AND METHOD

Differentiation of Energy with Respect to VL parameters

In the virtual ligand method, penalty functions corresponding to electronic/steric approximation are added to

the electronic energy.[9] Therefore, the potential energy surface (PES) is given as follows:

$$E(\boldsymbol{Q}, \boldsymbol{p}_{VL}, \boldsymbol{q}_{VL}) = E_{elec}(\boldsymbol{Q}) + V_{VL}(\boldsymbol{Q}, \boldsymbol{p}_{VL}, \boldsymbol{q}_{VL}),$$

where E_{elec} is the electronic energy determined by an atomic coordinate $\boldsymbol{Q}$, and V_{VL} is the penalty function determined by $\boldsymbol{Q}$, electronic/steric parameters $\boldsymbol{p}_{VL}$ and internal

parameters q_{VL} in the VL. Assuming that q_{VL} is optimized at every update of Q by microiteration (*vide infra*), we can describe the energy of an EQ as follows:

$$E_{EQ}(p_{VL}) = E(Q^*(p_{VL}), p_{VL}, q_{VL}^*(Q^*(p_{VL}), p_{VL})).$$

Here, $q_{VL}^*(Q, p_{VL})$ is the internal parameters minimizing E with given Q and p_{VL}, *i.e.*,

$$q_{VL}^*(Q, p_{VL}) = \underset{q_{VL}}{\mathrm{argmin}}[E(Q, p_{VL}, q_{VL})],$$

and $Q^*(p_{VL})$ is the local minimum of a function $Q \mapsto E(Q, p_{VL}, q_{VL}^*(Q, p_{VL}))$ in a neighborhood of an initial guess. The derivative of E_{EQ} with respect to VL parameters can be calculated as follows:

$$\begin{aligned}\frac{dE_{EQ}(p_{VL})}{dp_{VL}} = {} & \frac{\partial E}{\partial p_{VL}}(Q^*, p_{VL}, q_{VL}^*) \\ & + \frac{\partial E}{\partial Q}(Q^*, p_{VL}, q_{VL}^*)\frac{dQ^*}{dp_{VL}}(p_{VL}) \\ & + \frac{\partial E}{\partial q_{VL}}(Q^*, p_{VL}, q_{VL}^*)\frac{\partial q_{VL}^*}{\partial p_{VL}}(p_{VL}).\end{aligned}$$

At the local minima, the atomic coordinate Q^* and the internal parameter q_{VL}^* are optimized so that the first-order derivatives of the energy are zero, *i.e.*,

$$\frac{\partial E}{\partial Q}(Q^*, p_{VL}, q_{VL}^*) = 0,$$
$$\frac{\partial E}{\partial q_{VL}}(Q^*, p_{VL}, q_{VL}^*) = 0.$$

Therefore, the equation above can be rewritten as

$$\begin{aligned}\frac{dE_{EQ}(p_{VL})}{dp_{VL}} &= \frac{\partial E}{\partial p_{VL}}(Q^*, p_{VL}, q_{VL}^*) \\ &= \frac{\partial V_{VL}}{\partial p_{VL}}(Q^*, p_{VL}, q_{VL}^*).\end{aligned}$$

This equation claims that if the penalty function in the VL (V_{VL}) is differentiable, the energy of a local minima can be differentiated with respect to electronic/steric parameters, regardless of what approximation is used. Similarly, the energy of a TS can be differentiated as follows:

$$\frac{dE_{TS}(p_{VL})}{dp_{VL}} = \frac{\partial V_{VL}}{\partial p_{VL}}(Q^\circ, p_{VL}, q_{VL}^*),$$

where Q° is the atomic coordinate of the corresponding TS. It should be noted that, throughout this study, the electronic energies of EQs and TSs were used instead of the corresponding Gibbs free energies because of the difficulty in differentiating the thermal correction term. Although this could cause some errors in the computational prediction, we believe that this approximation does not affect the big picture of phenomena discussed in this paper, allowing us to draw at least qualitative conclusions.

Virtual Ligand Used in This Study

In this paper, we adopted a newly developed VL for all calculations. In this virtual ligand, a substituent is replaced with a Cl atom having some penalty functions (denoted as Cl*). For example, as previously reported,[9] a monodentate phosphorus(III) ligand PR_3 can be approximated using PCl^*_3 moiety. In addition, the new method is applicable for multidentate phosphorus(III) ligands. For example, as described later, 1,2-bis(diphenylphosphino)ethane (dppe) and its derivatives can be approximated by $Cl^*_2PCH_2CH_2PCl^*_2$ moiety.

The penalty functions for the Cl* atoms are formulated in the following part of this section.

Electronic Approximation

For approximation of electronic properties of phosphorus(III) ligands, we focused on the local Tolman's electronic parameter (LTEP)[13] and the pyramidalization parameter proposed by Gavrish.[14] In a VL, these properties of phosphine ligands were reproduced by two potential functions, the keep potential and the keep angle potential (Figure 2). The keep potential is a harmonic potential on P–Cl* distance and formulated as follows:

$$V_{keep} = \sum_{i \in Cl^*} \frac{1}{2}k_{keep}(r_i - r_0),$$

where r_i is a distance between i-th Cl* atom and phosphorus atom bonded to the Cl* atom. The k_{keep} and r_0 are the force constant and the equilibrium distance of the harmonic potential, respectively. The k_{keep} value was set to 4.48×10^3 kcal/ (mol Å²) (2.0 au) as previously reported.[9] On the other hand, the keep angle potential is a harmonic potential on the angle between a P–Cl* bond and a direction of the lone pair orbital, which is formulated as follows:

$$V_{keep\,angle} = \sum_{i \in Cl^*} \frac{1}{2}k_{keep\,angle}(\phi_i - \phi_0),$$

where ϕ_i is an angle between i-th P–Cl* bond and the direction of the lone pair orbital, which is defined based on the directions of three substituents on the phosphorous atom. The $k_{keep\,angle}$ and ϕ_0 are the force constant and the equilibrium angle of the harmonic potential, and the $k_{keep\,angle}$ value was set to 1.79×10^4 kcal/ (mol Å²) (8.0 au). By carefully tuning r_0 and ϕ_0 in a VL, both the LTEP and the pyramidalization parameter of real phosphorus(III) ligands can be reproduced (see Figure S1 for detail). It should be noted that, although the r_0 and ϕ_0 values to reproduce electronic effects of a given ligand slightly changes depends on the choices of the force constants (k_{keep} and $k_{keep\,angle}$), this would not affect the accuracy of the approximation as long as the LTEP and the pyramidalization parameter are reproduced.
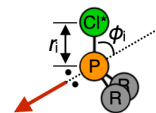
Figure 2. Definitions of r_i and ϕ_i in the keep and keep angle potentials for electronic approximation.

Steric Approximation

In the VL used in this study, the steric effect of a substituent is reproduced using the "ovoid-based" Lennard–Jones potential (hereafter, called as the ovoid LJ potential). The ovoid has seven geometrical parameters a_1, a_2, b_1, b_2, c_1, c_2 and d, which define size, shape, and position of the ovoid (Figure 3a). Size and shape of the ovoid are defined by stitching one-eighth surface of eight ellipsoids which have either a_1 or a_2, b_1 or b_2 and c_1 or c_2 as their axis (Figure 3b). The position of the ovoid is defined by the parameter d and the Cl* atom, where the center of the ovoid is located on the direction of the P–Cl* bond, and distance between P and the center of the ovoid is defined by d. By carefully selecting the

seven parameters, geometrical features of real substituents can be reproduced using the ovoid (Figure 3c, see Figure S2 and S3 for the detailed procedure and other examples).

In quantum chemical calculations, the steric effect of a substituent was reproduced by calculating the [12,6]-Lenard–Jones potential based on the ovoid. According to the description of nonbonded interactions in the universal force field (UFF) model,[15] the penalty value corresponding to the steric interaction between i-th ovoid and j-th counterpart (either an atom or an ovoid) was calculated as follows:

$$V_{LJ}^{ij} = \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right],$$

where r_{ij} is the distance between i-th ovoid and j-th counterpart. The ε_{ij} and σ_{ij} are the geometric means of the well depth or the van der Waals distance of i-th and j-th particles, respectively. For normal atoms, homonuclear parameters of the well depth and the van der Waals distance in the UFF scheme[15] were adopted. For the ovoid, these parameters were calculated as follows. Apparently, the van der Waals distance is correlated with the van der Waals radius. Assuming that the distance where the homonuclear LJ potential becomes zero is close to the diameter of the corresponding atom (Figure 4a), this correlation can be formulated as follows:

$$2r_{vdw} = \frac{1}{\sqrt[6]{2}} \sigma,$$

where r_{vdw} is the van der Waals radius, and σ is the homonuclear parameters of the van der Waals distance in the UFF scheme. Indeed, as shown in Figure 4b, this relationship was confirmed for non-metallic elements up to the third period.

Therefore, the van der Waals distance of the ovoid can be calculated based on its "radius". The "radius" of the ovoid (i) from the viewpoint of the counterpart (j) varies depending on their relative configuration (Figure 4c). Therefore, the intersection of the surface of the ovoid and the line connecting the centers of them was determined, and the van der Waals radius of the ovoid was calculated as the distance between the center of the ovoid and the intersection. Then, the van der Waals distance of the ovoid was calculated according to the above equation. The well depth of the ovoid was set to be 0.4 kcal/mol after numerically testing 0–1.0 kcal/mol (see Figure S4 for detail).

The overall penalty for steric interactions were calculated as the summation of penalties for every ovoid-atom and ovoid-ovoid pair. The formulation is as follows:

$$V_{LJ} = \sum_{i \in \mathbf{O}} \sum_{j \in \mathbf{S}} V_{LJ}^{ij} + \sum_{i \in \mathbf{O}} \sum_{j \in \mathbf{O}, j < i} V_{LJ}^{ij},$$

where $\mathbf{O}$ is a set of ovoids, and $\mathbf{S}$ is a set of normal atoms. Exceptionally, and as is the case in the UFF model,[15] ovoid-atom interactions for pairs that are bonded to each other (1,2 interactions) or bonded to a common atom (1,3 interactions) were assumed to be zero.

The overall penalty for steric interactions (V_{LJ}) varies depending on the configuration of the ovoids. Therefore, the rotations of the ovoids around P–Cl* axis, which are corresponding to the rotations of substituents, were considered as shown in Figure 4d. The rotation angles of all ovoids $\mathbf{q}_{VL} = (q_1, q_2, \ldots, q_M)$ were optimized so that the overall penalty (V_{LJ}) is minimized. The detailed algorithm for optimization and workflow of microiteration is described in Figure S5.
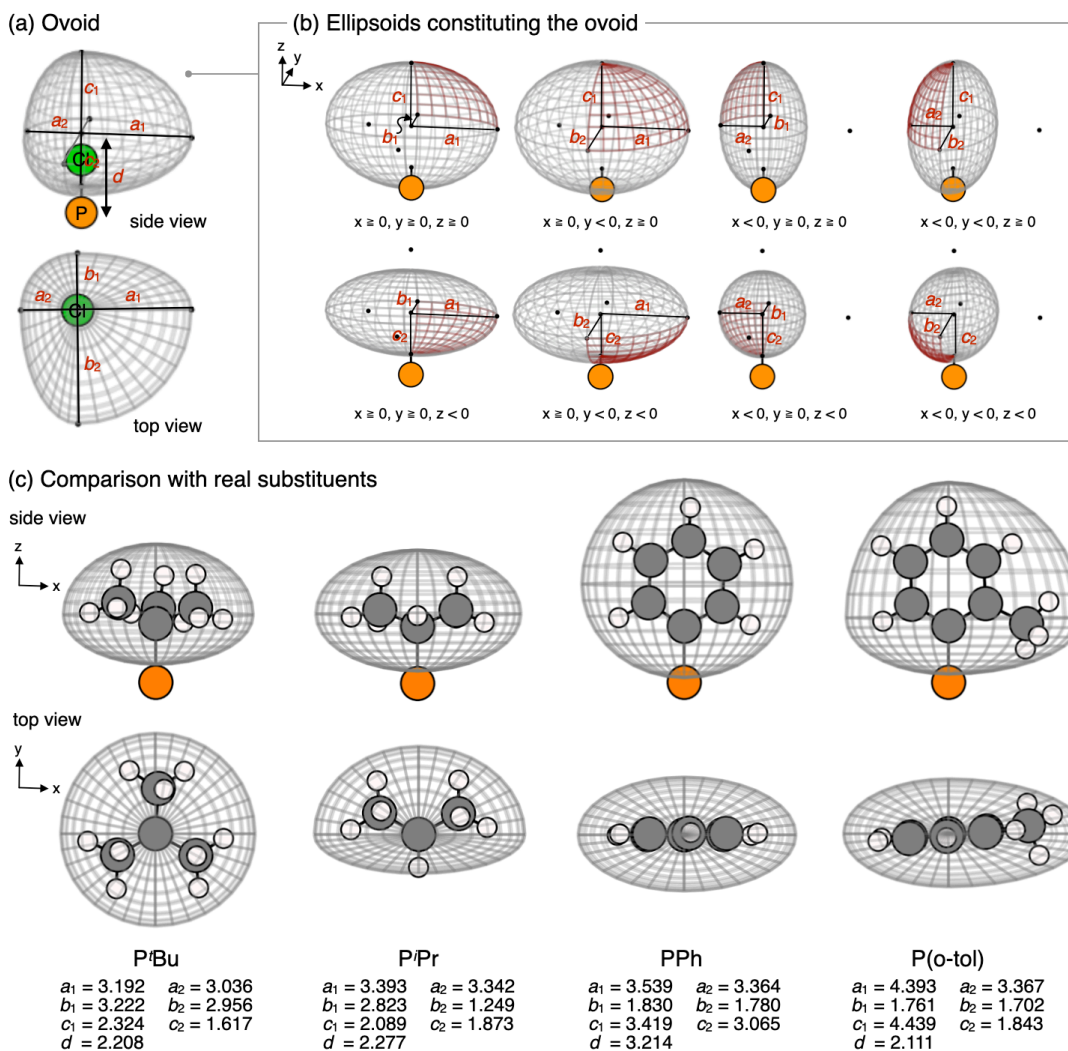
Figure 3. The ovoid used for the ovoid LJ potential. (a) Seven geometrical parameters which defines the size, shape, and position of the ovoid. (b) Eight ellipsoids constituting the ovoid. (c) Examples and comparisons of geometrical parameters for ^{t}Bu, ^{i}Pr, Ph and *o*-tol groups bonded to a phosphorous atom. In (c), all parameters are reported in angstroms.
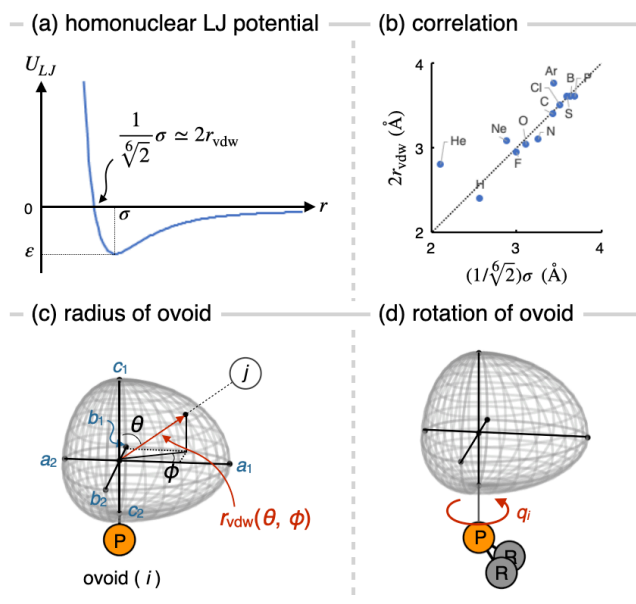
Figure 4. (a) Schematic illustration of the homonuclear LJ potential. (b) Correlation between r_{vdw} and σ. (c) Definition of the van der Waals radius of the ovoid. (d) Definition of an internal parameter q_i corresponding to the rotation of a substituent.

Calculation of Derivatives

All derivative values including the gradient, Hessian (with respect to the atomic coordinate Q) arisen from the penalty function (V_{LJ}), as well as derivatives of energy with respect to VL parameters (dE_i/dp_{VL}), were calculated using PyTorch module for Python.[16] In the Hessian calculation, the so-called effective Hessian[17] was calculated to take the quadratic couplings between an atomic coordinate Q and internal parameters q_{VL} into account.

RESULT AND DISCUSSION

Case Study 1: Optimization of Monodentate Phosphine

As a proof-of-concept study, we firstly focused on the Cu-catalyzed 1,3-dipolar cycloaddition of azomethine ylide **1** with methacrylonitrile **2** reported by Houk, Overman and coworkers (Figure 5a).[18] They reported that the choice of

phosphine ligands largely affects diastereoselectivity of the reaction. For example, whereas a 32:68 mixture of (*endo*)-**3** and (*exo*)-**3** was obtained when PPh$_3$ was employed, the selectivity was improved up to 10:90 by using PCy$_3$ as a ligand. They analyzed the origin of diastereoselectivity by DFT calculations and revealed that a steric repulsion between the nitrile group in **2** and PCy$_3$ destabilizes the TS for the *endo* cycloaddition, resulting in the selective formation of the *exo* cycloadduct.

To check if the VLAO method can reproduce this ligand engineering, we performed quantum chemical calculations using the VL. First, the VL parameters were tuned to reproduce PPh$_3$ (hereafter, this parameter set is referred as $\boldsymbol{p}_{VL}{}^{Ph}$): Electronic parameters r_0 and ϕ_0 were tuned to be 1.606 Å and 64.215°, respectively, to reproduce both LTEP and pyramidalization parameter of PPh$_3$ (2174.5 cm^{-1} and 7.14 degree$^{1/2}$). Steric parameters in the ovoid LJ-potential were set to reproduce the steric effect of PPh$_3$, where a_1, a_2, b_1, b_2, c_1, c_2 and d were tuned to fit geometric feature of the phenyl group bonded to a phosphorous atom as shown in Figure 3c. Then, geometry optimizations of transition states for the 1,3-dipolar cycloaddition were performed using thus-tuned VL, and (*endo*)-**TS** was calculated to be more stable than (*exo*)-**TS** by 0.44 kcal/mol (Figure 5b). The *exo* selectivity (ξ_{exo}) was estimated based on the following equation:

$$\xi_{exo} = \frac{\exp(-\frac{E_{exo}^{\ddagger}}{RT})}{\exp(-\frac{E_{exo}^{\ddagger}}{RT}) + \exp(-\frac{E_{endo}^{\ddagger}}{RT})} \times 100,$$

where R and T are the gas constant and the temperature. The temperature was set to be 300.0 K. The $E_{exo}^{\ddagger}$ and $E_{endo}^{\ddagger}$ are the electronic energies of (*exo*)-**TS** and (*endo*)-**TS**, respectively. According to this equation, the *exo* selectivity was calculated to be 32.5%, which is slightly underestimated comparing to the experimental result using PPh$_3$ (68%). With the structure of (*endo*)-**TS** and (*exo*)-**TS** in hand, we then performed the differentiation of the electronic energy, and the derivative of the *exo* selectivity in terms of VL parameters was calculated (see SI for the detailed formulation). The derivative values with respect to parameters a_1, a_2, b_1, b_2, c_1, c_2 and r_0 were shown in Figure 5b (right). The positive derivative values for steric parameters (a_1, a_2, b_1, b_2, c_1 and c_2) imply that the *exo* selectivity can be improved by increasing the steric bulk of the substituent. Judging from the magnitude of derivative values, it turned out that the parameters for thickness (b_1, b_2) and length (c_2) of the ovoid has larger impacts on the *exo* selectivity. As for the electronic parameter (r_0), a small negative derivative was obtained. Because an increase of r_0 value makes the VL electron-deficient (see Figure S1), this result implies that electron-donating ligands would increase the *exo* selectivity. Because Cy group is apparently thicker and more electron-donating than Ph group, these results are qualitatively consistent with ligand engineering by Overman *et. al.*, where the *exo* selectivity was improved by the use of PCy$_3$.

Then, we have performed the optimization of VL parameters (Figure 5c). In this case, the objective function (F) was set to be the *exo* selectivity (ξ_{exo}) corrected by an additional penalty function (P) to constrain the ranges of parameters (see below), and seven VL parameters (r_0, a_1, a_2, b_1, b_2, c_1 and c_2) were optimized to maximize F. The parameter set which reproduce PPh$_3$ ($\boldsymbol{p}_{VL}{}^{Ph}$) was used as the initial parameter. The penalty function P constrains VL parameters to be optimized within "realistic" ranges. For example, the range of electronic parameter r_0 was set to be 1.2–2.4 Å, which covers the electronic effect between P^{t}Bu$_3$ and P(CF$_3$)$_3$. When the r_0 value approaches to the lower or upper limit, a positive penalty was subtracted from the objective function because real ligand molecules with such a high electron-donating or electron-withdrawing ability are rarely found (see Figure S6 for the detailed formulation of P). Similarly, the ranges of steric parameters were set to be 1.0–6.0 Å for a_1, a_2, b_1, b_2 and c_1, and 1.0–3.514 Å for c_2. The upper limit of c_2 (3.514 Å) was set considering the steric parameter d (3.214 Å). The parameter optimization was performed by the CG method. At each iteration, the VL parameters were updated based on the CG calculated using the derivative value of F ($dF/d\boldsymbol{p}_{VL}$), and geometry optimizations of (*exo*)-**TS** and (*endo*)-**TS** were performed. The *exo* selectivity (ξ_{exo}) as well as the objective function (F) were then computed and plotted in Figure 5c (crosses and open circles, respectively) along with corresponding VL parameters (dotted lines). After each line search, the maximum absolute values among the gradient components and the displacement components were calculated, and the convergence was determined when each of these values was less than 1 Å^{-1} and 0.1 Å, respectively. In addition, the objective function (F) was monitored at each iteration, and the calculation was terminated when the F value exceeds 99%. This criterion was set to prevent VL parameters from vibrating in a subspace where the F value is close to its upper limit (100%). As a result, the calculation was converged at after 17 iterations, and the F value was improved from 32.5% to 99.7%. The ovoid in the VL at the first and last iterations (ITR. 0 and ITR. 17) were visualized on Figure 5c (middle). While the electronic feature (expressed by the color of the ovoid) has been slightly changed, the thickness of the ovoid was increased significantly (about 0.8 Å in total). To check how this steric change affects the *exo* selectivity, steric interactions between each ovoid in the VLs and methacrylonitrile **2** were analyzed. In (*endo*)-**TS**, two repulsive interactions (**Int-1** and **Int-2**) were found: **Int-1** is the repulsion between the hydrogen at the terminal olefin (H36) and the ovoid 1, and **Int-2** is that between the nitrogen at the nitrile group (N42) and the ovoid 2. As summarized in the table, these repulsive interactions increased as the VL became bulkier. In particular, **Int-2** increased by 0.32 kcal/mol before and after optimization (0.05 kcal/mol at ITR. 0 and 0.37 kcal/mol at ITR. 17). Similar two repulsive interactions were also found in (*exo*)-**TS**: One is the repulsion of the hydrogen at the terminal olefin, and the other is that of the hydrogen at the methyl group. However, these repulsive interactions did not increase as much as **Int-2** did in the (*endo*)-**TS** when the VL became bulkier. Indeed, the repulsion of the hydrogen at the terminal olefin or the methyl group were calculated to be 0.14 and 0.06 kcal/mol even after optimization of VL parameters (ITR. 17). Thus, the steric change in the VL effectively destabilized only (*endo*)-**TS**, resulting in the increase of *exo* selectivity. This result is fully consistent with the ligand engineering by Houk and Overman *et. al.*[18]
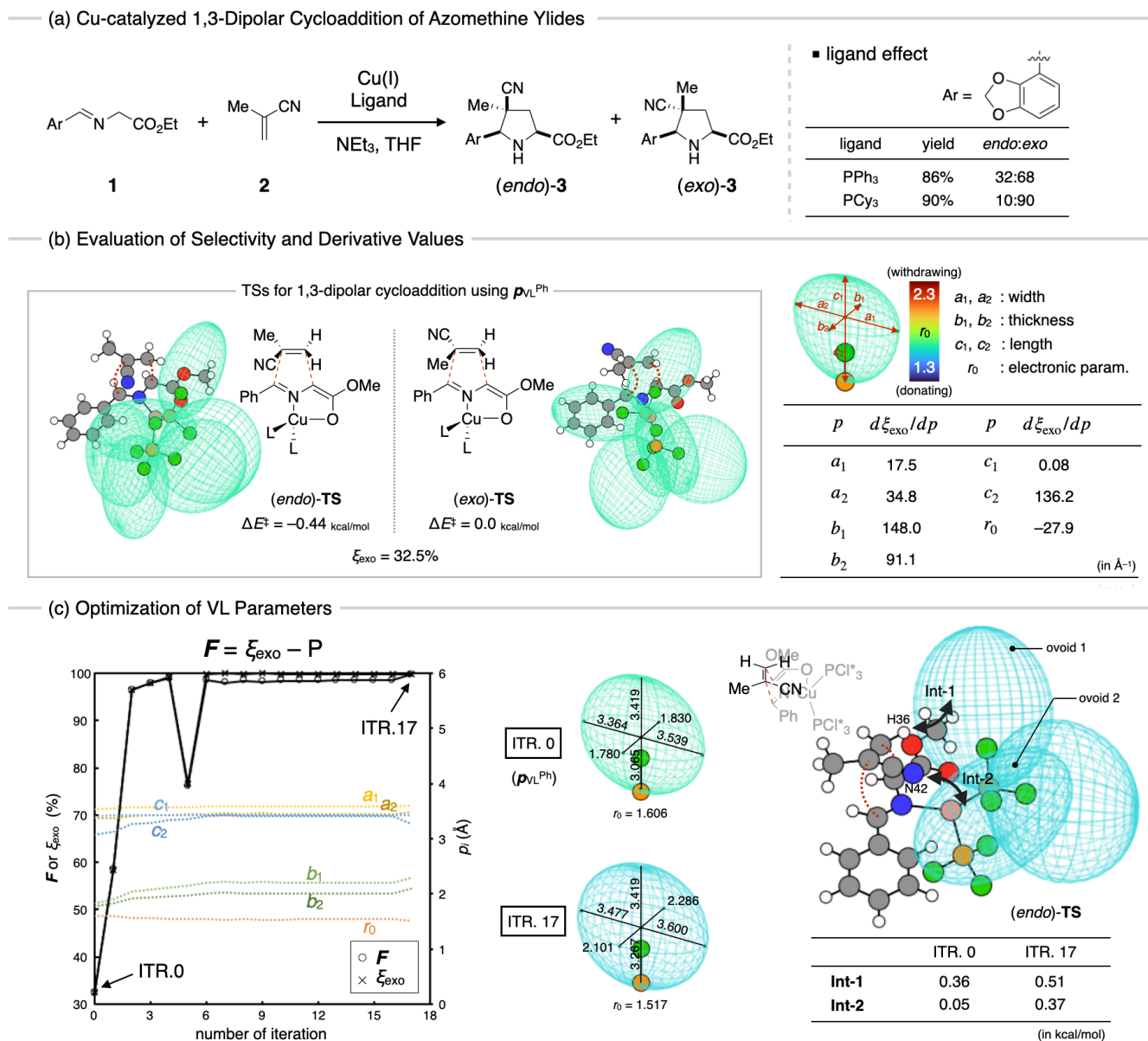
(a) Cu-catalyzed 1,3-Dipolar Cycloaddition of Azomethine Ylides

ligand	yield	*endo:exo*
PPh$_3$	86%	32:68
PCy$_3$	90%	10:90

(b) Evaluation of Selectivity and Derivative Values

TSs for 1,3-dipolar cycloaddition using $p_{VL}{}^{Ph}$

(endo)-**TS**
$\Delta E^{\ddagger} = -0.44$ kcal/mol

(exo)-**TS**
$\Delta E^{\ddagger} = 0.0$ kcal/mol

$\xi_{exo} = 32.5\%$

(withdrawing) / (donating)

a_1, a_2 : width
b_1, b_2 : thickness
c_1, c_2 : length
r_0 : electronic param.

p	$d\xi_{exo}/dp$	p	$d\xi_{exo}/dp$
a_1	17.5	c_1	0.08
a_2	34.8	c_2	136.2
b_1	148.0	r_0	−27.9
b_2	91.1		(in Å^{-1})

(c) Optimization of VL Parameters

$F = \xi_{exo} - P$

ITR. 0 ($p_{VL}{}^{Ph}$)
3.364 3.419 1.830 1.780 3.539 3.085
$r_0 = 1.606$

ITR. 17
3.477 3.419 2.286 2.101 3.600 3.267
$r_0 = 1.517$

(endo)-**TS**

	ITR. 0	ITR. 17
Int-1	0.36	0.51
Int-2	0.05	0.37
		(in kcal/mol)

Figure 5. The VLAO-based engineering of monodentate phosphine ligands. (a) General scheme of the Cu-catalyzed 1,3-dipolar cycloaddition of azomethine ylide **1** to methacrylonitrile **2**. (b) Evaluation of the *exo* selectivity and its derivative values using the VL tuned to reproduce PPh$_3$. (c) Optimization of the VL parameters to increase the objective function *F*. In right panel of (c), the ovoids corresponding to one of the VLs are omitted for clarity.

Case Study 2: Optimization of dppe-type Ligand

The VLAO method can be applied not only monodentate but also multidentate phosphine ligands. As a proof-of-concept study, we focused on the reductive elimination of trifluorotoluene from a palladium(II) diphosphine complex **5** (Figure 6a).[19] In general, this reaction is very challenging, and diphosphine ligands with a large bite angle, such as Xantphos, had been thought to be necessary to promote the desired reductive elimination. Therefore, 1,2-bis(diphenylphosphino)ethane and its analogues (dppe-type ligands), which have relatively small bite angles, had been thought to be unsuitable for this reaction. However, upon a sophisticated ligand engineering based on quantum chemical calculations, Schoenebeck and coworkers have reported that the desired reductive elimination can be effectively accelerated

by the dfmpe ligand, which is a dppe-type ligand with CF$_3$ substituents.[19] Indeed, the activation energy for the reaction using dfmpe ligand was calculated to be 25.4 kcal/mol (at the ωB97XD/Def2-SVP level), which is much smaller than that calculated using dppe (35.4 kcal/mol). The key principles of their ligand design were 1) the low steric demand of substituents on the phosphorous atoms, 2) large repulsion of the ligand with the Ph and CF$_3$ groups on the Pd(II) center, and 3) highly electron-withdrawing property of the ligand. The second and third principles are in line with the generally accepted trends for reductive elimination, whereas the first criterion is less intuitive because bulky ligands were considered to accelerate reductive elimination. Also, it should be noted that the first and second criteria generally contradict each other because the first criterion requires small substituents while the second one does large

substituents. Schoenebeck and coworkers elegantly met both requirements by introducing a small but highly polarized CF_3 group into the ligand.

We performed quantum chemical calculations using the VL to check if this ligand engineering can be reproduced by the VLAO method (Figure 6b). In this case, to approximate dppe-type ligands by a virtual ligand, 1,2-bis(dichlorophosphino)ethane was used as the dummy moiety, and the penalty functions (the keep potential, the keep angle potential and the ovoid LJ potential) were added to four Cl atoms. Firstly, to reproduce dppe with thus-prepared virtual ligand, geometry optimizations of the palladium(II) complex **4** and **TS$_{45}$** were performed using the VL parameters tuned to reproduce PPh$_3$ ($\boldsymbol{p}_{VL}^{Ph}$). As a result, the activation energy ($\Delta E^{\ddagger}$) was calculated to be 36.1 kcal/mol which is consistent with the value calculated using the actual dppe molecule (Figure 6a, 35.4 kcal/mol). We then performed the differentiation of the electronic energy of **4** and **TS$_{45}$**, and the derivative values of the activation energy ($\Delta E^{\ddagger}$) was calculated (see SI for the detailed formulation). The derivative values with respect to parameters a_1, a_2, b_1, b_2, c_1, c_2 and r_0 were shown in Figure 6b (right). As for the electronic parameter (r_0), a large negative derivative value was obtained. Because an increase of r_0 value makes the VL electron-deficient (see Figure S1), this result is consistent with the third criterion for the ligand design by Schoenebeck and coworkers (see above). Regarding the steric parameters (a_1, a_2, b_1, b_2, c_1 and c_2), small negative derivative values were computed, implying that increase of steric bulk would slightly decrease the activation energy. Because the first and second criteria for the ligand design would generally demand opposite steric features (see above), the small absolute values for steric parameters might be a result of the cancelling out of these irreconcilable effects.

Then, we performed the optimization of VL parameters. In this case, the objective function (F) was set to be the sum of the activation energy ($\Delta E^{\ddagger}$) and the range confining penalty function (P), and seven VL parameters (r_0, a_1, a_2, b_1, b_2, c_1 and c_2) were optimized to minimize F. As in the previous example, the parameter set which reproduce dppe ($\boldsymbol{p}_{VL}^{Ph}$) was used as the initial parameter, and the ranges of VL parameters were set to be 1.2–2.4 Å for r_0, 1.0–6.0 Å for a_1, a_2, b_1, b_2 and c_1, and 1.0–3.514 Å for c_2. The results are summarized in the bottom of Figure 6b. The calculation converged after nine iterations, and steric parameters (a_1, a_2, b_1, b_2, c_1 and c_2) changed little. Instead, the electronic parameter r_0 changed significantly, converging at 2.345 Å, which is close to the upper limit (2.4 Å). The activation energy ($\Delta E^{\ddagger}$) with the optimized parameter was calculated to be 14.9 kcal/mol, which is much lower than the calculated values for dfmpe or 1,2-bis(difluorophosphino)ethane (Figure 6a). This result could indicate that, if dppe-type ligands with even more electron-deficient substituents are synthesized, the desired reductive elimination can be further accelerated. Although designing such a highly electron-deficient molecule is not an easy task, this might be achieved by introducing unique design strategies for electron-withdrawing ligands, such as the use of cationic substituents[20] or replacing phosphorous atoms with arsenic[21].
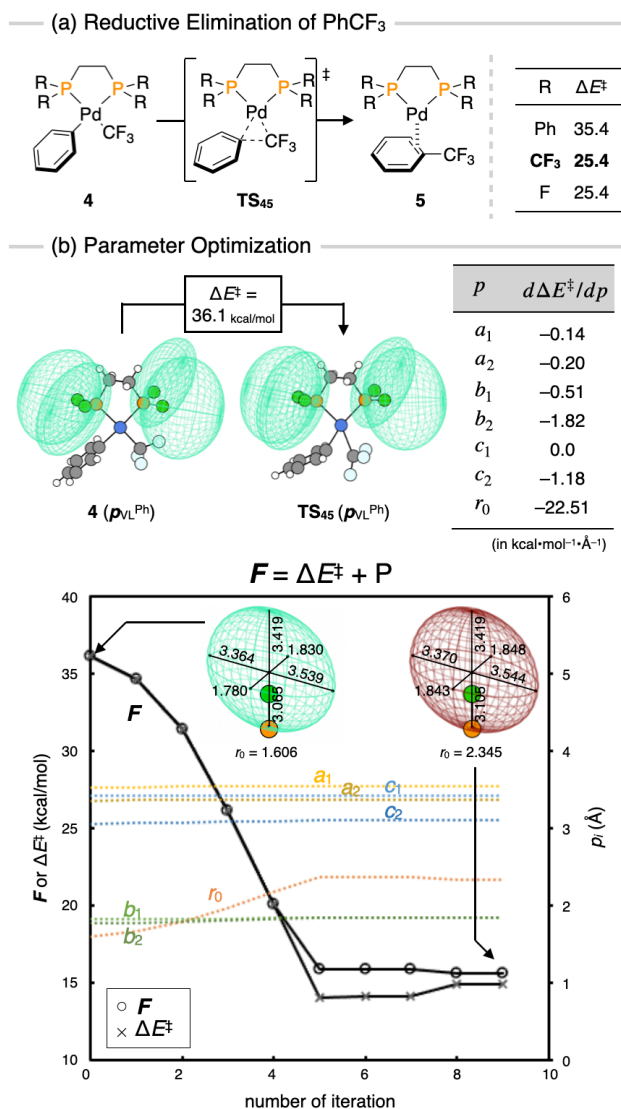
Figure 6. The VLAO-based engineering of dppe-type ligands. (a) General scheme of the reductive elimination of PhCF$_3$ from Pd(II) complex **4**. (b) Evaluation of the activation energy and its derivative values (top) and optimization of the VL parameters to minimize F (bottom).

Case Study 3: Optimization of Asymmetric Ligand

In the two examples above, all substituents in a VL were described by common electronic and steric parameters, thereby enabling design of symmetric phosphine ligands such as PR$_3$ and R$_2$PCH$_2$CH$_2$PR$_2$. Because electronic and steric parameters can be set independently for each substituent, the VLAO strategy would also be applicable to engineering of asymmetric phosphine ligands. To prove this, we focused on the Cu-catalyzed asymmetric hydroboration of aliphatic terminal alkenes reported by Ito and coworkers (Figure 7a).[22] The hydroboration of terminal alkene **6** potentially affords three isomeric products, namely (S)-**7**, (R)-**7** and **8**. Ito and coworkers found that, this reaction can be catalyzed with high regio- and enantioselectivity when (S)-quinox-^{t}Bu$_3$, a C_1-symmetric quinoxaline-based diphosphine ligands, was used. Then, the authors conducted detailed analyses of origin of observed selectivities based on both computational and experimental methods and

established a design principle: Whereas the regioselectivity is affected by R^2 and R^4, the enantioselectivity is determined by R^2 and R^3 (Figure 7a, right). According to this guideline, the authors then designed and synthesized an unprecedented asymmetric ligand, namely (S)-quinox-^{t}OctAd$_2$. Indeed, this ligand enabled the hydroboration of **6** in excellent regio- and enantioselectivity (92 and 99%, respectively).

To reproduce this ligand engineering with the VLAO method, we calculated eight TSs of the borylcupration according to Ito's computational analysis (Figure 7b). These TSs correspond to eight reaction routes of borylcupration, where an olefine (1-butene) approaches to the Cu–B bond from any of the four directions (denoted by **A**, **B**, **C** or **D**) to give either the branched product **7** or the linear product **8** (denoted by subscript **b** or **l**). Out of the four routes to give the branched product, **A$_b$** and **D$_b$** give the major enantiomer (S)-**7**, whereas **B$_b$** and **C$_b$** afford the minor enantiomer (R)-**7**. The regio- and enantioselectivity of the reaction are determined by the relative energies of these TSs, and the branch selectivity (bs) and enantiomeric excess (ee) can be estimated as follows:

$$\text{bs} = \frac{\sum_{i=A_b,B_b,C_b,D_b} \exp(-\frac{E_i^{\ddagger}}{RT})}{\sum_{i=A_b,B_b,C_b,D_b,A_l,B_l,C_l,D_l} \exp(-\frac{E_i^{\ddagger}}{RT})} \times 100,$$

$$\text{ee} = \frac{\sum_{i=A_b,D_b} \exp(-\frac{E_i^{\ddagger}}{RT}) - \sum_{i=B_b,C_b} \exp(-\frac{E_i^{\ddagger}}{RT})}{\sum_{i=A_b,D_b} \exp(-\frac{E_i^{\ddagger}}{RT}) + \sum_{i=B_b,C_b} \exp(-\frac{E_i^{\ddagger}}{RT})} \times 100,$$

where R and T are the gas constant and the temperature, respectively. The temperature was set to be 233.15 K. The $E_i^{\ddagger}$ is the electronic energy of a TS corresponding to the route i. In this case, the VL with the 2,3-bis(dichlorophosphino)quinoxaline moiety was used to reproduce quinoxaline-based diphosphine ligands. As shown in Figure 7b (middle), the electronic and steric parameters of each substituent were set independently, with those of the Me group ($p_{VL}{}^{Me}$) for R^1 and those of the ^{t}Bu group ($p_{VL}{}^{tBu}$) for R^2–R^4, to reproduce (S)-quinox-^{t}Bu$_3$ as a whole. It should be noted that, considering high symmetry of Me and ^{t}Bu groups as well as the potential difficulty in optimizing many independent parameters, steric parameters corresponding to width and thickness of each substituent were set to be symmetric (i.e., $a_1 = a_2$ and $b_1 = b_2$ in each substituent). Geometry optimizations of eight TSs using thus-prepared VL were performed, and the branch selectivity and the enantiomeric excess were calculated to be 97.9% and 82.9%, respectively. These values are reasonably consistent with the experimental results (bs = 72% and ee = 87%) though the branch selectivity is slightly overestimated. With the optimized structures of eight TSs in hand, we then calculated the derivative values of the branch selectivity and the enantiomeric excess in terms of VL parameters (see SI for the detailed formulation). The derivative values with respect to steric parameters corresponding to width and thickness of each ovoid were shown in Figure 7b (right). In this table, the a_{1-i} and b_{1-i} represent the width and thickness parameters for the i-th ovoid (R^i). As for the derivative values of R^1 (a_{1-1} and b_{1-1}), whose effect was not discussed in Ito's report, intuitive results were obtained: An increasing steric bulk of R^1 increases the branch selectivity and decreases of the

enantiomeric excess. If the Me group in (S)-quinox-^{t}Bu$_3$ is replaced with a bulkier substituent, steric repulsion between the ligand and the alkyl chain in the olefin would increase, thereby resulting in the selective formation of the branched product through routes **A$_b$**, **B$_b$**, **C$_b$** or **D$_b$**. On the other hand, because the Me group is the only substituent which breaks symmetry of the ligand, replacing it with larger substituents would make the whole ligand (pseudo) C_2 symmetric, resulting in the decreased enantioselectivity. Regarding the effects of R^2–R^4 (a_{1-2}, b_{1-2}, a_{1-3}, b_{1-3}, a_{1-4} and b_{1-4}) on the enantioselectivity ($d$ee$/dp$), the calculate derivative values were consistent with the trend reported by Ito et. al.: The enantioselectivity is mainly determined by the size of R^2 and R^3, and contribution of R^3 is greater than that of R^2 (see the magnitude of the derivative values). On the other hand, the effect of these substituents on the regioselectivity was relatively small comparing to those on the enantioselectivity and was not entirely consistent with the reported trends. Whereas Ito and coworkers reported that the regioselectivity is affected by the bulk of R^2 and R^4, the derivative values ($d$bs$/dp$) imply that not only these substituents but also R^3 have contribution to the regioselectivity. To see the origin of this contribution, steric interactions between the ovoid corresponding to R^3 and substrates (i.e., 1-butene and Bpin) in the TS for the route **A$_l$**, which is the most favorable reaction path to give the linear product **8**, were analyzed (Figure 7c, left). As a result, it was found that the strongest interaction is the repulsion between the ovoid and the hydrogen atom at the olefin (H26). Thus, increasing bulk of R^3 (a_{1-3} or b_{1-3}) would enhance this repulsive interaction, destabilizing the corresponding TS. This will in turn cause the selective formation of the branched product, resulting in the positive derivative values for a_{1-3} and b_{1-3}. Actually, this interaction was also found in the optimized structure for the corresponding TS reported by Ito and coworkers. The shortest interatomic distance between the hydrogen atom and the ^{t}Bu group corresponding to R^3 was calculated to be 2.34 Å (Figure 7c, right). This value is comparable to the shortest interatomic distances between the Bpin moiety and the ^{t}Bu groups corresponding to R^2 or R^4, on which the importance of R^2 and R^4 was discussed based. Because C_1-symmetric quinox-type ligands which have different substituents at R^3 and R^4 positions are synthetically unavailable, the importance of R^3 could have been masked by that of R^4 and overlooked. Hence, the VLAO approach which enables systematic analyses of ligand effects can be a powerful tool for identifying unapparent but valuable factors for ligand engineering. To evaluate overall importance of each substituent, the derivative values of the sum of the branch selectivity and enantioselectivity were calculated (d(bs + ee)$/dp$). Based on this result, criteria for each substituent to improve both regio- and enantioselectivity are suggested: R^1 should be smaller than Me group, R^2 and R^3 should be larger than ^{t}Bu group, and R^4 has the smallest impact on the selectivity. As Me group is the smallest alkyl group, and R^4 must be the same substituent as R^3 (see above), these criteria would be consistent with the ligand engineering by Ito and coworkers, where (S)-quinox-^{t}OctAd$_2$ was identified as the optimal ligand, considering synthetic accessibility of ligands.

Finally, the optimization of the VL parameters were performed (Figure 7d). The objective function (F) was set to be

the sum of the branch selectivity (bs) and enantiomeric excess (ee) corrected by the range confining penalty function (P). The parameter set which reproduce (S)-quinox-^{t}Bu$_3$ were used as the initial guess, and eight steric parameters (a_{1-1}, b_{1-1}, a_{1-2}, b_{1-2}, a_{1-3}, b_{1-3}, a_{1-4} and b_{1-4}) were optimized. The ranges were set to be 1.0–6.0 Å for all parameters. The results are summarized in Figure 7d, in which the objective function (F), the branch selectivity (bs) and the enantiomeric excess (ee) are plotted with dots, crosses and open circles, respectively. As expected from the derivative values at the initial step (Figure 7b), bulk of the ovoids corresponding to R^2 and R^3 (a_{1-2}, b_{1-2}, a_{1-3} and b_{1-3}) increased significantly, while the parameters for R^4 remained almost constant. On the other hand, despite the negative derivative values at the initial step (Figure 7b), bulk of

the ovoid corresponding to R^1 (a_{1-1} and b_{1-1}) converged at slightly bigger values than those for Me group. This is probably because, as optimization proceeds, R^1 became no longer the only substituent which breaks symmetry of the ligand unlike it was at the first step of the optimization (see above), allowing R^1 to be a larger substituent. Overall, the optimization terminated after 13 iterations, where the branch selectivity and the enantiomeric excess were estimated to be extremely high (99.9% and 99.4%, respectively). As the optimized VL parameters demand substituents R^1–R^4 to have different size and shape to each other, real ligand molecules which meet these criteria would be out of the current synthetic scope. However, we believe that the VLAO method can accelerate discovery of useful ligands by motivating chemists to explore such untapped chemical spaces.
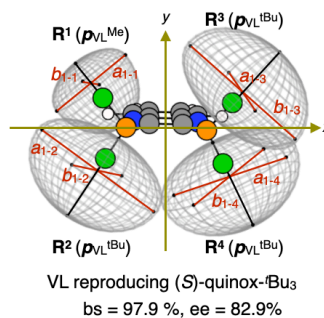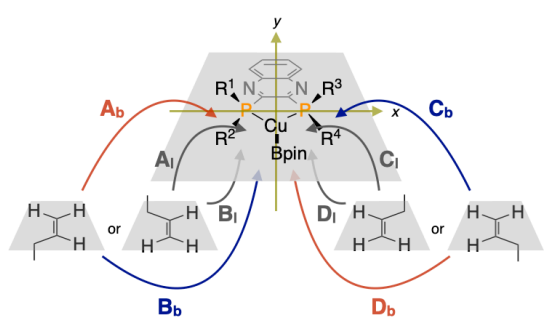

(a) Asymmetric Hydroboration of Aliphatic Terminal Alkene
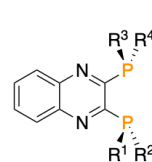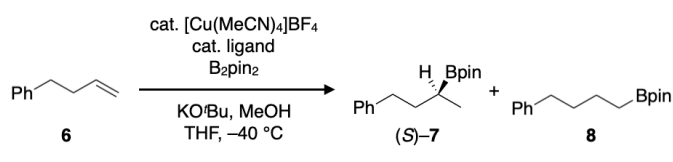
ligand	R^1	R^2	R^3, R^4	bs	ee
(S)-quinox-^{t}Bu$_3$	Me	^{t}Bu	^{t}Bu	72%	87%
(S)-quinox-^{t}OctAd$_2$	Me	^{t}Oct	1-Ad	92%	99%

regioselectivity: affected by R^2, R^4
enantioselectivity: affected by R^2, R^3

(b) Selectivity-Determining TS and Ligand Effect

p	$d\,\mathrm{bs}/dp$	$d\,\mathrm{ee}/dp$	$d(\mathrm{bs+ee})/dp$
a_{1-1}	3.65	−4.64	−0.99
b_{1-1}	2.33	−10.80	−8.47
a_{1-2}	0.43	1.42	1.85
b_{1-2}	2.95	12.09	15.05
a_{1-3}	0.57	8.02	8.58
b_{1-3}	2.58	20.64	23.22
a_{1-4}	0.55	−0.57	−0.03
b_{1-4}	1.68	−3.62	−1.94

(in Å^{-1})

VL reproducing (S)-quinox-^{t}Bu$_3$
bs = 97.9 %, ee = 82.9%

(c) Finding of important Interactions

0.23 kcal/mol (repulsion)

with VL

2.34 Å
2.28 Å   2.37 Å

with (S)-quinox-^{t}Bu$_3$ (reported by Ito *et. al.*)

TS for route **A$_l$** (the most favorable route to give **8**)

(d) Parameter Optimization

$$F = \mathrm{bs} + \mathrm{ee} - P$$
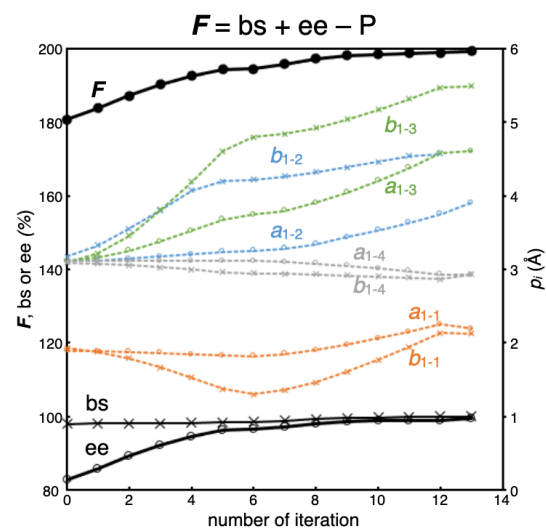
Figure 7. The VLAO-based engineering of asymmetric ligands. (a) General scheme of the Cu-catalyzed asymmetric hydroboration of terminal olefin. (b) Evaluation of branch selectivity (bs), enantiomeric excess (ee) and their derivative values. (c) Finding of an important interaction to improve regioselectivity. (d) Optimization of the VL parameters to maximize F.

Experimental Validation

With the potential utility of the VLAO method proved by reproducing the reported ligand engineering processes, we then performed the experimental validation of the VLAO

method. For this purpose, we have focused on the palladium-catalyzed hydrogermylation of ynamide **9** reported by Blanchard, Houk, Bizet and coworkers.[23] The hydrogermylation of **9** affords two regioisomeric products, namely α-product **10** and β-product **11**, and the

regioselectivity can be controlled by ligands (Figure 8a). For example, hydrogermylation of **9a** using bis[2-(diphenylphosphino)phenyl] ether (DPEPhos) affords the α-product selectively, whereas the β-product is exclusively formed when PMe^{t}Bu$_2$ is employed as the ligand. On the contrary, hydrogermylation of **9b**, which has a terminal alkyne moiety, tends to give the β-product preferentially, and, to the best of our knowledge, no α-selective ligands have been reported. Hence, we aimed to find α-selective ligands by the VLAO method. First, using the virtual ligand reproducing PPh$_3$, selectivity-determining TSs, namely **α-TS** and **β-TS**, were calculated, and the α selectivity ξ_α was estimated to be 40.7%. according to the following equation:

$$\xi_\alpha = \frac{\exp(-\frac{E_\alpha^\ddagger}{RT})}{\exp(-\frac{E_\alpha^\ddagger}{RT}) + \exp(-\frac{E_\beta^\ddagger}{RT})} \times 100,$$

where $E_\alpha^\ddagger$ and $E_\beta^\ddagger$ are the electronic energies of **α-TS** and **β-TS**, respectively (Figure 8b). The temperature was set to be 333.15 K. Then, the ξ_α were differentiated with respect to the electronic and steric parameters (see SI for the detailed formulation), and optimization of these parameters based on the CG method was performed. Again, the α selectivity corrected by the range confining penalty function (P) was used as the objective function, and the optimization was terminated when this value exceeded 99%. As a result, the calculation was terminated after 17 iterations, and the initial and resulting ovoids were depicted in Figure 8b (right). While b_1 parameter slightly increased (2.050 Å), b_2 and c_2 parameters decreased (1.377 Å and 2.707 Å, respectively), resulting in a thinner and shorter ovoid than the initial one. As for electronic effect, r_0 parameter decreased significantly, converging at 1.301 Å, which is close to the lower limit (1.2 Å). Overall, the VLAO method implied that the α selectivity can be largely improved by the use of thinner and shorter substituents with stronger electron-donating ability comparing to Ph group.

With the suggestion by the VLAO in hand, we then aimed at finding α-selective ligands (Figure 8c). Considering the requirement for electron-donating ligands, we prepared six phosphine ligands including PMe$_3$, PEt$_3$, PCy$_3$, PCp$_3$ (Cp = cyclopentyl), **L1** and **L2**. Whereas trialkylphosphine ligands (PMe$_3$, PEt$_3$, PCy$_3$ and PCp$_3$) were chosen because they are well known to exhibit strong electron-donating nature, **L1** and **L2** were designed so that they have stronger electron-donating ability than PPh$_3$ while keeping the steric effect unchanged as much as possible. The α selectivities for each ligand were then predicted. It

should be noted that, whereas some of real ligands (*e.g.*, PCy$_3$ and PCp$_3$) are very flexible, the VL parameters for a given ligand were determined based on a single conformer of R$_3$PNi(CO)$_3$ (see SI for detail). Hence, the VL only reproduces a snapshot of a ligand and cannot describe conformational flexibility (except for rotation of P–R bonds). In this study, to avoid this inherent limitation of the VL method, the prediction of the α selectivity for a given ligand was performed as follows. First, upon a systematical conformation search by the SC-AFIR method[24], the most stable conformer of R$_3$PNi(CO)$_3$ were determined. Then, all conformers with relative energies lower than 1.0 kcal/mol were chosen, and VL parameters for each of these conformers were determined (see Table S2– S5). After that, the calculations of **α-TS** and **β-TS** using each set of VL parameters were performed, and the α selectivity of the ligand was estimated as the average value of those for each conformer weighted by the Boltzmann distribution of R$_3$PNi(CO)$_3$. As a result, the α selectivities were estimated to be 88.4% for PMe$_3$, 36.5% for PEt$_3$, 4.2% for PCy$_3$, 29.5% for PCp$_3$, 54.1% for **L1** and 98.7% for **L2**, respectively. As expected, both **L1** and **L2** were predicted to exhibit higher α selectivity than PPh$_3$ (40.7%). On the contrary, the α selectivities of PEt$_3$, PCy$_3$ and PCp$_3$ were predicted to be lower than PPh$_3$ despite their high electron-donating ability. This is because these alkyl groups are much thicker than the phenyl group, making corresponding ligands sterically unfavorable. It is important to noted that the α selectivity of PCp$_3$ (29.5%) differs significantly from the value estimated from only the most stable conformer (98.5%). This result highlights the importance of the contributions from not only the most stable but also thermally accessible conformers of a ligand.

Then, the experiments using these ligands as well as PPh$_3$ were performed, and the observed yields and selectivities were summarized in the table (Figure 8c). To our surprise, PPh$_3$ exhibit higher α selectivity than DPEPhos, which was determined as the optimal ligand for α-selective hydrogermylation of **9a** (Figure 8a). The observed selectivities were plotted against predicted values (Figure 8c), and a moderate correlation for all ligands were confirmed ($R^2 = 0.59$). Overall, the highest observed α selectivity was 73% for **L2**, and this value this value did not change when the catalyst loading was reduced to 1.5 mol%. Notably, **L2** is less commercially available compared to other typical triarylphosphines probably because it is rapidly oxidized when the solution is exposed to air. Hence, it is unlikely to be included as an initial candidate for random screening of ligands. The VLAO method rapidly suggested important features of ligand, enabling us to find such an elusive ligand with a minimum number of experimental trials.
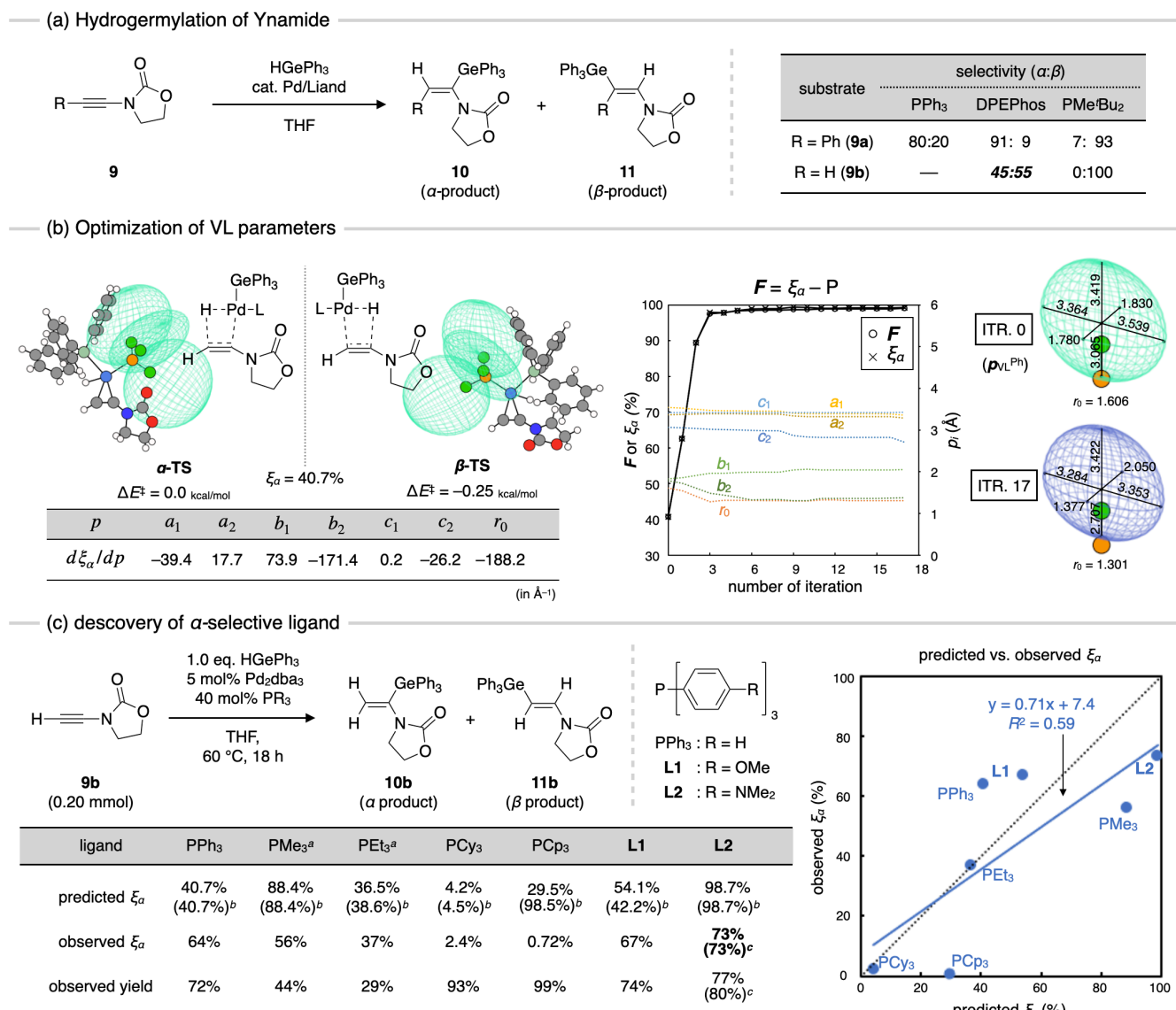
(a) Hydrogermylation of Ynamide

substrate	selectivity (α:β)		
	PPh$_3$	DPEPhos	PMe^{t}Bu$_2$
R = Ph (9a)	80:20	91: 9	7: 93
R = H (9b)	—	**45:55**	0:100

9 → 10 (α-product) + 11 (β-product)

HGePh$_3$, cat. Pd/Liand, THF

(b) Optimization of VL parameters

α-TS, $\Delta E^{\ddagger}$ = 0.0 kcal/mol, ξ_α = 40.7%
β-TS, $\Delta E^{\ddagger}$ = −0.25 kcal/mol

p	a_1	a_2	b_1	b_2	c_1	c_2	r_0
$d\xi_\alpha/dp$	−39.4	17.7	73.9	−171.4	0.2	−26.2	−188.2

(in Å^{-1})

$F = \xi_\alpha - P$

ITR. 0 (p_{VL}^{Ph}): 3.364, 3.419, 1.830, 1.780, 3.065, 3.539; r_0 = 1.606

ITR. 17: 3.284, 3.422, 2.050, 1.377, 2.795, 3.353; r_0 = 1.301

(c) discovery of α-selective ligand

9b (0.20 mmol) → 10b (α product) + 11b (β product)

1.0 eq. HGePh$_3$, 5 mol% Pd$_2$dba$_3$, 40 mol% PR$_3$, THF, 60 °C, 18 h

PPh$_3$: R = H
L1 : R = OMe
L2 : R = NMe$_2$

ligand	PPh$_3$	PMe$_3^a$	PEt$_3^a$	PCy$_3$	PCp$_3$	L1	L2
predicted ξ_α	40.7% (40.7%)^b	88.4% (88.4%)^b	36.5% (38.6%)^b	4.2% (4.5%)^b	29.5% (98.5%)^b	54.1% (42.2%)^b	98.7% (98.7%)^b
observed ξ_α	64%	56%	37%	2.4%	0.72%	67%	**73% (73%)^c**
observed yield	72%	44%	29%	93%	99%	74%	77% (80%)^c

predicted vs. observed ξ_α

$y = 0.71x + 7.4$
$R^2 = 0.59$

Figure 8. Experimental validation of the VLAO-based ligand engineering. (a) General scheme of the Pd-catalyzed hydrogermylation of ynamides. (b) Evaluation of the α selectivity (ξ_α) and optimization of VL parameters. (c) Discovery of α-selective ligand. ^{a}PR$_3$•HBF$_4$ (40 mol%) and Cs$_2$CO$_3$ (100 mol%) were used instead of PR$_3$. ^{b}Calculated using VL parameters derived from the most stable conformation of R$_3$PNi(CO)$_3$. ^{c}1.5 mol% Pd$_2$dba$_3$ and 12 mol% **L2** were used.

CONCLUSION

In conclusion, we have reported the virtual ligand-assisted optimization (VLAO) method, a rational strategy for ligand engineering. In this method, the electronic and steric effects of phosphine ligands on the reaction outcome can be mathematically quantified by differentiating energies of equilibrium structures and transition states in terms of electronic and steric parameters in the virtual ligand. The obtained derivative values can be used as a rational principle of ligand optimization. Indeed, the VLAO method reproduced reported design for monodentate, bidentate and asymmetric phosphine ligands. In addition, an optimal ligand for the α-selective hydrogermylation of terminal ynamide **9b** was rapidly designed and discovered based on the guiding principle by the VLAO method.

On the other hand, because of a large number of parameters in the virtual ligand, it is not a simple task to transfer the optimized parameters to a real molecule. Therefore, the discussions about ligand engineering in this paper were only qualitative. In addition, as showcased by the example of PCp$_3$ (Figure 8c), because a single set of VL parameters cannot describe conformational flexibility of a ligand, the prediction must be made carefully (*e.g.*, use the Boltzmann weighted average). We believe that these drawbacks can be overcome once a machine learning model that maps VL parameters to the chemical space of ligands while considering conformational flexibility is established. The optimized parameters can be then utilized as quantitative guides to pick up promising candidates from databases[25] such as the ligand knowledge base (LKB) and the kraken. Execution of such informatic processing is one of the tasks in the near future.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Computational procedures, computational data, experimental procedures, experimental data, and additional discussion (PDF)

Cartesian coordinates of optimized structures (xyz)

AUTHOR INFORMATION

Corresponding Author

* matsuoka.wataru@sci.hokudai.ac.jp (W.M.)

* iwata@mist.i.u-tokyo.ac.jp (S. I.)

* smaeda@eis.hokudai.ac.jp (S.M.)

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

REFERENCES

1. For selected reviews focusing on specific classes of organic ligands, see: (a) Noyori, R.; Takaya, H. BINAP: An Efficient Chiral Element for Asymmetric Catalysis. *Acc. Chem. Res.* **1990**, *23*, 345–350. (b) Hayashi, T. Chiral Monodentate Phosphine Ligand MOP for Transition-Metal-Catalyzed Asymmetric Reactions. *Acc. Chem. Res.* **2000**, *33*, 354–362. (c) Kamer, P. C. J.; van Leeuwen, P. W. N. M.; Reek, J. N. H. Wide Bite Angle Diphosphines: Xantphos Ligands in Transition Metal Complexes and Catalysis. *Acc. Chem. Res.* **2001**, *34*, 895–904. (d) Gómez Arrayás, R.; Adrio, J.; Carretero, J. C. Recent Applications of Chiral Ferrocene Ligands in Asymmetric Catalysis. *Angew. Chem., Int. Ed.* **2006**, *45*, 7674–7715. (e) Martin, R.; Buchwald, S. L. Palladium-Catalyzed Suzuki–Miyaura Cross-Coupling Reactions Employing Dialkylbiaryl Phosphine Ligands. *Acc. Chem. Res.* **2008**, *41*, 1461–1473. (f) Ye, B.; Cramer, N. Chiral Cyclopentadienyls: Enabling Ligands for Asymmetric Rh(III)-Catalyzed C–H Functionalizations. *Acc. Chem. Res.* **2015**, *48*, 1308–1318. (g) Janssen-Müller, D.; Schlepphorst, C.; Glorius, F. Privileged Chiral *N*-Heterocyclic Carbene Ligands for Asymmetric Transition-Metal Catalysis. *Chem. Soc. Rev.* **2017**, *46*, 4845–4854. (h) Zhao, Q.; Meng, G.; Nolan, S. P.; Szostak, M. *N*-Heterocyclic Carbene Complexes in C–H Activation Reactions. *Chem. Rev.* **2020**, *120*, 1981–2048.

2. For selected revies focusing on ligand design, see: (a) Berthod, M.; Mignani, G.; Woodward, G.; Lemaire, M. Modified BINAP: The How and the Why. *Chem. Rev.* **2005**, *105*, 1801–1836. (b) 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. (c) Engle, K. M.; Yu, J.-Q. Developing Ligands for Palladium(II)-Catalyzed C–H Functionalization: Intimate Dialogue between Ligand and Substrate. *J. Org. Chem.* **2013**, *78*, 8927–8955. (d) Peris, E.; Crabtree, R. H. Key Factors in Pincer Ligand Design. *Chem. Soc. Rev.* **2018**, *47*, 1959–1968. (e) Fanourakis, A.; Docherty, P. J.; Chuentragool, P.; Phipps, R. J. Recent Developments in Enantioselective Transition Metal Catalysis Featuring Attractive Noncovalent Interactions between Ligand and Substrate. *ACS Catal.* **2020**, *10*, 10672–10714.

3. Tolman, C. A. Steric Effects of Phosphorus Ligands in Organometallic Chemistry and Homogeneous Catalysis. *Chem. Rev.* **1977**, *77*, 313–348.

4. For a selected review focusing of ligand descriptors, see: Durand, D. J.; Fey, N. Computational Ligand Descriptors for Catalyst Design. *Chem. Rev.* **2019**, *119*, 6561–6594.

5. For selected examples of electronic parameters, 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) Suresh, C. H.; Koga, N. Quantifying the Electronic Effect of Substituted Phosphine Ligands via Molecular Electrostatic Potential. *Inorg. Chem.* **2002**, *41*, 1573–1578. (e) Coll, D. S.; Vidal, A. B.; Rodríguez, J. A.; Ocando-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. (f) 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.

6. For selected examples of steric parameters, 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*. (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. (c) Bilbrey, 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*, 2741–2770.

7. For selected revies focusing on ligand design by informatic techniques, see: (a) Zahrt, A. F.; Athavale, S. V.; Denmark, S. E. Quantitative Structure–Selectivity Relationships in Enantioselective Catalysis: Past, Present, and Future. *Chem. Rev.* **2020**, *120*, 1620–1689. (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. (c) Duan, C.; Nandy, A.; Kulik, H. J. Machine Learning for the Discovery, Design, and Engineering of Materials. *Annu. Rev. Chem. Biomol. Eng.* **2022**, *13*, 405–429.

8. For selected examples of ligand engineering by informatic techniques, see: (a) Harper, K. C.; Bess, E. N.; Sigman, M. S. Multidimensional Steric Parameters in the Analysis of Asymmetric Catalytic Reactions. *Nat. Chem.* **2012**, *4*, 366–374. (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) Newman-Stonebraker, S. H.; Smith, S. R.; Borowski, J. E.; Peters, E.; Gensch, T.; Johnson, H. C.; Sigman, M. S.; Doyle, A. G.

Univariate Classification of Phosphine Ligation State and Reactivity in Cross-Coupling Catalysis. *Science* **2021**, *374*, 301–308. (f) Akana, M. E.; Tcyrulnikov, S.; Akana-Schneider, B. D.; Reyes, G. P.; Monfette, S.; Sigman, M. S.; Hansen, E. C.; Weix, D. J. Computational Methods Enable the Prediction of Improved Catalysts for Nickel-Catalyzed Cross-Electrophile Coupling. *J. Am. Chem. Soc.* **2024**, *146*, 3043–3051.

9. (a) Matsuoka, W.; Harabuchi, Y.; Maeda, S. Virtual Ligand-Assisted Screening Strategy to Discover Enabling Ligands for Transition Metal Catalysis. *ACS Catal.* **2022**, *12*, 3752–3766. (b) Matsuoka, W.; Harabuchi, Y.; Nagata, Y.; Maeda, S. Highly Chemoselective Ligands for Suzuki–Miyaura Cross-Coupling Reaction Based on Virtual Ligand-Assisted Screening. *Org. Biomol. Chem.* **2023**, *21*, 3132–3142. (c) Matsuoka, W.; Harabuchi, Y.; Maeda, S. Virtual Ligand Strategy in Transition Metal Catalysis Toward Highly Efficient Elucidation of Reaction Mechanisms and Computational Catalyst Design. *ACS Catal.* **2023**, *13*, 5697–5711.

10. (a) Koga, N.; Morokuma, K. A Simple Scheme of Estimating Substitution or Substituent Effects in the Ab Initio MO Method Based on the Shift Operator. *Chem. Phys. Lett.* **1990**, *172*, 243–248. (b) 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. (c) Kozuch, S.; Shaik, S. Defining the Optimal Inductive and Steric Requirements for a Cross-Coupling Catalyst Using the Energetic Span Model. *J. Mol. Catal. A: Chem.* **2010**, *324*, 120–126.

11. (a) Fernández-Ramos, A.; Miller, J. A.; Klippenstein, S. J.; Truhlar, D. G. Modeling the Kinetics of Bimolecular Reactions. *Chem. Rev.* **2006**, *106*, 4518–4584. (b) Kozuch, S.; Shaik, S. How to Conceptualize Catalytic Cycles? The Energetic Span Model. *Acc. Chem. Res.* **2011**, *44*, 101–110. (c) Rush, L. E.; Pringle, P. G.; Harvey, J. N. Computational Kinetics of Cobalt-Catalyzed Alkene Hydroformylation. *Angew. Chem., Int. Ed.* **2014**, *53*, 8672‑8676. (d) Klippenstein, S. J.; Pande, V. S.; Truhlar, D. G. Chemical Kinetics and Mechanisms of Complex Systems: A Perspective on Recent Theoretical Advances. *J. Am. Chem. Soc.* **2014**, *136*, 528–546.

12. (a) Sumiya, Y.; Nagahata, Y.; Komatsuzaki, T.; Taketsugu, T.; Maeda, S. Kinetic Analysis for the Multistep Profiles of Organic Reactions: Significance of the Conformational Entropy on the Rate Constants of the Claisen Rearrangement. *J. Phys. Chem. A* **2015**, *119*, 11641–11649. (b) Sumiya, Y.; Maeda, S. A Reaction Path Network for Wöhler's Urea Synthesis. *Chem. Lett.* **2019**, *48*, 47–50. (c) Sumiya, Y.; Maeda, S. Rate Constant Matrix Contraction Method for Systematic Analysis of Reaction Path Networks. *Chem. Lett.* **2020**, *49*, 553–564. (d)Iwata, S.; Oki, T.; Sakaue, S. Rate Constant Matrix Contraction Method for Stiff Master Equations with Detailed Balance. *arXiv* **2023**, 2312.05470 [math.NA].

13. Kalescky, R.; Kraka, E.; Cremer, D. New Approach to Tolman's Electronic Parameter Based on Local Vibrational Modes. *Inorg. Chem.* **2014**, *53*, 478–495.

14. Gavrish, S. P. Approximate Expressions and the Relationship between Pyramidalization (Out-of-plane Deformation) Characteristics of Trigonal Centers. *J. Comput. Chem.* **2012**, *33*, 2173–2179.

15. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.

16. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Kopf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; Chintala, S. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems* **2019**, *32*, 8024–8035.

17. Maeda, S.; Ohno, K.; Morokuma, K. An Automated and Systematic Transition Structure Explorer in Large Flexible Molecular Systems Based on Combined Global Reaction Route Mapping and Microiteration Methods. *J. Chem. Theory Comput.* **2009**, *5*, 2734–2743.

18. Walton, M. C.; Yang, Y.-F.; Hong, X.; Houk, K. N.; Overman, L. E. Ligand-Controlled Diastereoselective 1,3-Dipolar Cycloadditions of Azomethine Ylides with Methacrylonitrile. *Org. Lett.* **2015**, *17*, 6166–6169.

19. (a) Nielsen, M. C.; Bonney, K. J.; Schoenebeck, F. Computational Ligand Design for the Reductive Elimination of $ArCF_3$ from a Small Bite Angle Pd^{II} Complex: Remarkable Effect of a Perfluoroalkyl Phosphine. *Angew. Chem., Int. Ed.* **2014**, *53*, 5903–5906. (b) Anstaett, P.; Schoenebeck, F. Reductive Elimination of $ArCF_3$ from Bidentate Pd^{II} Complexes: A Computational Study. *Chem. Eur. J.* **2011**, *17*, 12340–12346.

20. (a) Andrieu, J.; Azouri, M.; Richard, P. First Donor Stabilized-Phosphenium Rhodium Complexes. *Inorg. Chem. Commun.* **2008**, *11*, 1401–1404. (b) Petuškova, J.; Bruns, H.; Alcarazo, M. Cyclopropenylylidene-Stabilized Diaryl and Dialkyl Phosphenium Cations: Applications in Homogeneous Gold Catalysis. *Angew. Chem., Int. Ed.* **2011**, *50*, 3799‑3802. (c) Carreras, J.; Gopakumar, G.; Gu, L.; Gimeno, A.; Linowski, P.; Petuškova, J.; Thiel, W.; Alcarazo, M. Polycationic Ligands in Gold Catalysis: Synthesis and Applications of Extremely π-Acidic Catalysts. *J. Am. Chem. Soc.* **2013**, *135*, 18815–18823. (d) Tinnermann, H.; Nicholls, L. D. M.; Johannsen, T.; Wille, C.; Golz, C.; Goddard, R.; Alcarazo, M. *N*-Arylpyridiniophosphines: Synthesis, Structure, and Applications in Au(I) Catalysis. *ACS Catal.* **2018**, *8*, 10457–10463.

21. (a) Farina, V.; Krishnan, B. Large Rate Accelerations in the Stille Reaction with Tri-2-Furylphosphine and Triphenylarsine as Palladium Ligands: Mechanistic and Synthetic Implications. *J. Am. Chem. Soc.* **1991**, *113*, 9585–9595. (b) Tanaka, S.; Konishi, M.; Imoto, H.; Nakamura, Y.;Ishida, M.; Furuta, H.; Naka, K. Fundamental Study on Arsenic(III) Halides (AsX_3 ; X = Br, I) toward the Construction of C_3-Symmetrical Monodentate Arsenic Ligands. *Inorg. Chem.* **2020**, *59*, 9587–9593. (c) Chishiro, A.; Konishi, M.; Inaba, R.; Yumura, T.; Imoto, H.; Naka, K. Tertiary Arsine Ligands for the Stille Coupling Reaction. *Dalton Trans.* **2022**, *51*, 95–103.

22. Iwamoto, H.; Imamoto, T.; Ito, H. Computational Design of High-Performance Ligand for Enantioselective Markovnikov Hydroboration of Aliphatic Terminal Alkenes. *Nat. Commun.* **2018**, *9*, 2290.

23. Debrauwer, V.; Turlik, A.; Rummler, L.; Prescimone, A.; Blanchard, N.; Houk, K. N.; Bizet, V. Ligand-Controlled Regiodivergent Palladium-Catalyzed Hydrogermylation of Ynamides. *J. Am. Chem. Soc.* **2020**, *142*, 11153–11164.

24. 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.

25. For selected examples of phosphine databases, see: (a) 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. Euro. J.* **2006**, *12*, 291–302. (b) 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. (c) 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. (d) Gugler, S.; Janet, J. P.; Kulik,

H. J. Enumeration of de Novo Inorganic Complexes for Chemical Discovery and Machine Learning. *Mol. Syst. Des. Eng.* **2020**, *5*, 139–152. (e) Balcells, D.; Skjelstad, B. B. tmQM Dataset–Quantum Geometries and Properties of 86k Transition Metal Complexes. *J. Chem. Inf. Model.* **2020**, *60*, 6135–6146. (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. *J. Am. Chem. Soc.* **2022**, *144*, 1205–1217. (g) Chen, S.-S.; Meyer, Z.; Jensen, B.; Kraus, A.; Lambert, A.; Ess, D. H. ReaLigands: A Ligand Library Cultivated from Experiment and Intended for Molecular Computational Catalyst Design. *J. Chem. Inf. Model.* **2023**, *63*, 7412–7422.

Insert Table of Contents artwork here

Virtual Ligand-Assisted Optimization (**VLAO**)

$$\frac{dF}{d\boldsymbol{p}_{\mathrm{VL}}} = \Sigma \frac{\partial F}{\partial E_i} \frac{dE_i}{d\boldsymbol{p}_{\mathrm{VL}}}$$

design principle

a new strategy for in silico ligand engineering