Statistical Analysis of Catalytic Performance in Ethylene/Methyl Acrylate Copolymerization Using Palladium/Phosphine‑Sulfonate Catalysts

Shumpei Akita,[†] *Jing-Yao Guo,*[‡] *Falk W. Seidel, Matthew S. Sigman,**[‡] *Kyoko Nozaki**[†]

[†]Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

[‡]Department of Chemistry, College of Science, The University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112, United States

ABSTRACT: For various types of palladium complexes bearing phosphine-sulfonate (PS) ligands used in the coordination–insertion copolymerization of olefins with polar monomers, characteristic features of the ligands, such as electronic and steric properties, have been discussed to describe their catalytic performance. Aiming at further analysis of the literature data, here we report the development of a statistical method for how the ligand impacts the performance of a Pd-catalyzed copolymerization of ethylene and methyl acrylate. During our investigation, ligand features important for the resultant molecular weight of the obtained polymers were identified. Consistent with previously suggested important parameters, the electron density on the palladium center and

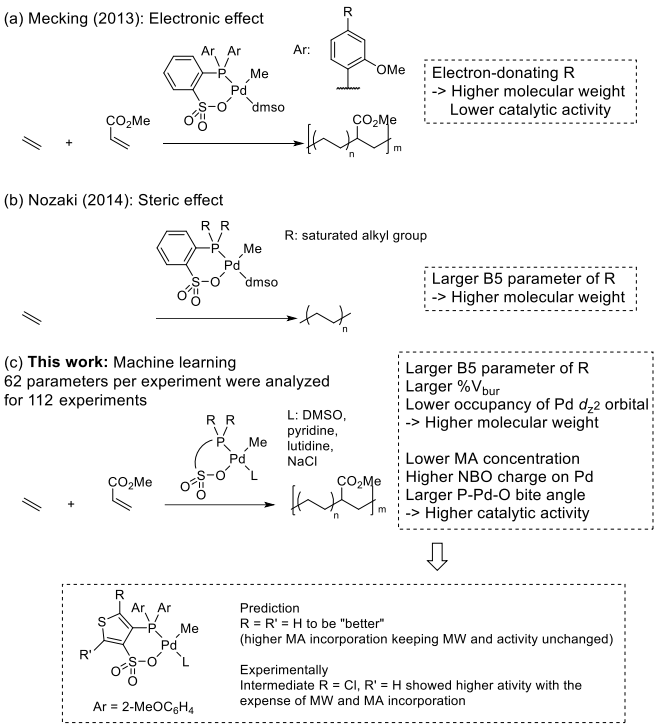
maximum width of the substituents on the phosphorus atom (B5) were found to be significant for catalyst performance. We also found that additional features impact reaction outputs. As an example, the lower occupancy of the palladium $d_{z}2$ orbital results in increase of molecular weight and catalyst activity in both ethylene homopolymerization and ethylene/methyl acrylate copolymerization. Furthermore, it was predicted that a larger bite angle of the ligand increased activity of ethylene/methyl acrylate copolymerization without impacting the molecular weight. On the basis of these machine learning predictions, three thiophene derived PS-type catalysts were synthesized and tested for MA/ethylene copolymerization. Unexpectedly, rather than the one predicted to enhance catalytic performance, a synthetic intermediate to this ligand exhibited higher activity albeit with the expense of molecular weight and MA incorporation. The inconsistency between the prediction and the experimental result is likely a result of insufficient training data for the catalyst with different linker moiety. However, the unexpected finding that chlorination of the ligand backbone increases the overall catalyst performance will inspire an avenue for PS catalyst development.

Introduction

Since the publication of the seminal work by Drent *et al.* in 2002[1], extensive effort has been devoted to the development of unsymmetrical bidentate ligands effective in promoting the palladium-catalyzed copolymerization of olefins with polar vinyl monomers.[2] Among the potent ligand types, phosphine–sulfonates (PS), which selectively facilitate the production of linear copolymers in high molecular weight (MW), have been the most investigated.[3] Various strategies have been developed over the years for the optimization of their catalytic performance, such as

activity, product MW, and increased incorporation of polar monomers, and the impact of ligand characteristics has been scrutinized with considerable effort. Notably, Wucher *et al.* reported the ligand electronic effect in the copolymerization of ethylene with methyl acrylate (MA) promoted by di-*o*-anisylphosphino-benzenesulfonates (4-**R**-2-MeO-C_6H_3)$_2$P-C_6H_4-SO_3), where the presence of a strong electron-donor on the phosphine component was observed to produce copolymers with increased MW, but with a concomitant decrease in the catalytic activity (**Scheme 1a**).[4] In 2014, Nozaki and coworkers examined the steric impact of the phosphine substituents with the feature Sterimol B5, which describes the maximum width of the substituent.[3a, 5] In both the homopolymerization of ethylene, and its copolymerization with a variety of polar monomers, the B5 measure is positively correlated with the resultant MW (**Scheme 1b**). This association was postulated to be caused by the ability of bulky substituents to effectively shield the apical position of the metal center. To further evaluate the interplay of ligand electronic and steric features on their catalytic properties, we envisaged applying more sophisticated data science tools to comprehensively investigate the stereoelectronic factors that control the performance of this catalyst class.

Scheme 1. Copolymerization of ethylene and methyl acrylate (MA) with palladium phosphine-sulfonate catalysts.



The product MW of a polymerization process is determined by the difference between the energies of activation ($\Delta\Delta G^{\ddagger}$) of chain-propagation and chain-transfer. Unfortunately, computational transition state studies with modern theoretical approaches (*e.g.*, DFT calculations) are not adequate for the assessment of energy differences on this scale. For example, a difference of 1.0 kcal/mol, which is approximately the limit of accuracy in a DFT calculation for transition-metal catalysis, corresponds to an increase in MW by ~4 fold.[6,7] Alternatively, the utilization of various molecular featurization tools, combined with machine learning algorithms, has been proven effective in the prediction of enantiomeric excess, a similarly sensitive reaction output.[8] For instance, in 2019, Denmark and co-workers demonstrated their process through the successful description of catalyst structural effects on the enantioselective thiol addition to *N*-acylimines, promoted by BINOL-derived phosphoric acids. In this study, the development of a steric descriptor that represents the entire conformer ensemble was the key to their success. In the same year,

Sigman and co-workers reported the statistical modelling of data-mined literature results from a range of nucleophilic addition reactions to imines using DFT-derived physical organic parameters, and the ability of the resulting model to provide out-of-sample predictions. We became interested in applying these types of data science methods to interrogate the performance of the Pd/PS system in ethylene polymerization and ethylene/MA copolymerization. Herein, we report the implementation of these types of workflows to explore which key catalyst features impact various polymerization metrics. Moreover, based on the statistical models constructed, we designed, synthesized and tested three new thiophene derived PS-type catalysts. In contrast to our prediction, the designed thiophene-linker PS catalyst exhibited only modest activity while a synthetic intermediate in route to this catalyst provides some of the best catalyst performance metrics reported to date.

Results and Discussion

Strategy

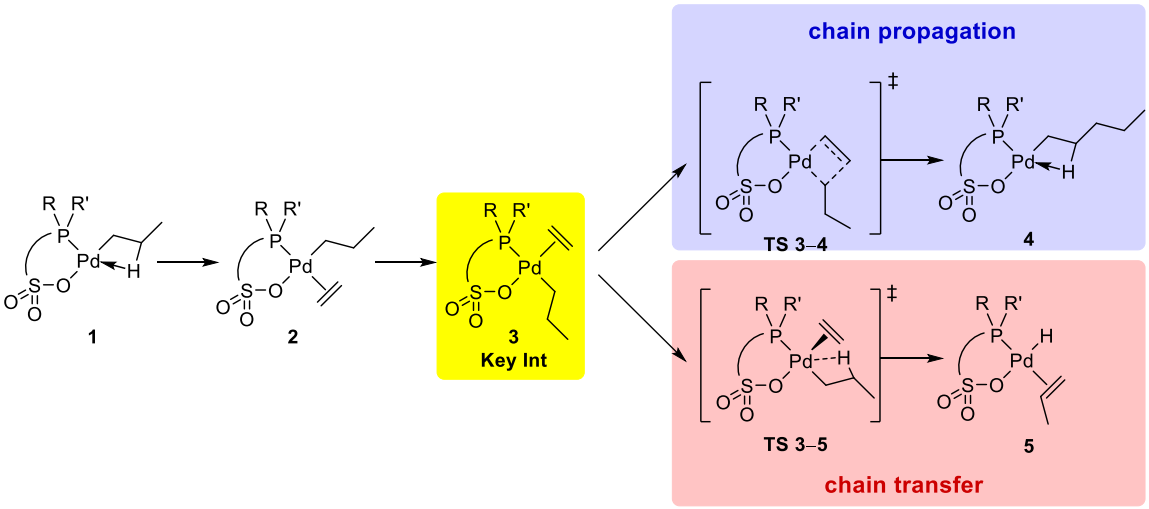
To initiate the investigation, we considered the general reaction mechanisms for chain-propagation and chain-transfer as suggested by previous studies (**Scheme 2**). DFT calculations suggested that the chain-propagation proceeds via the ethylene coordination to an alkylpalladium intermediate (**1**→**2**), followed by cis/trans isomerization (**2**→**3**), and ethylene insertion (**3**→**4**), with the final one (**3**→**4**) being the rate determining step of the desired polymerization reaction.[9] The rate-determining step for the chain-transfer, which limits the polymer-chain length, is ethylene dissociation from intermediate **3** to give **5** via β-hydride elimination. This mechanism suggests that intermediate **3** is the key intermediate located at the branching point to control molecular weight of a polymer. Accordingly, we selected this general structure to build the requisite *in silico*

feature set for statistical modelling applications.[10] The set of structures assessed in this study is tabulated in **Figure 1**.[3,9,11]

Comparative molecular field analysis (CoMFA) [11] was first employed to provide a spatial visualization of the catalyst properties and their effect on MW in ethylene homopolymerization. Subsequently, a statistical modelling study was carried out to determine how catalyst features impact (co)polymer properties (MW, polymerization activity, incorporation of polar monomer in case of E/MA copolymerization) in both homopolymerization of ethylene (E) and copolymerization of ethylene and methyl acrylate (MA).

In this paper, the polar monomer for copolymerization is limited to methyl acrylate because of biased the reported data sets. In 120 entries, methyl acylate was used as polar monomer although there are 188 examples for copolymerization by Pd/PS. Original references for the data are listed in Tables S1 and S2 in Supplementary Information.

Scheme 2. Mechanism of ethylene polymerization promoted by palladium/phosphine–sulfonate catalysts: Intermediate **3** is located at the branching point for chain propagation and chain transfer.



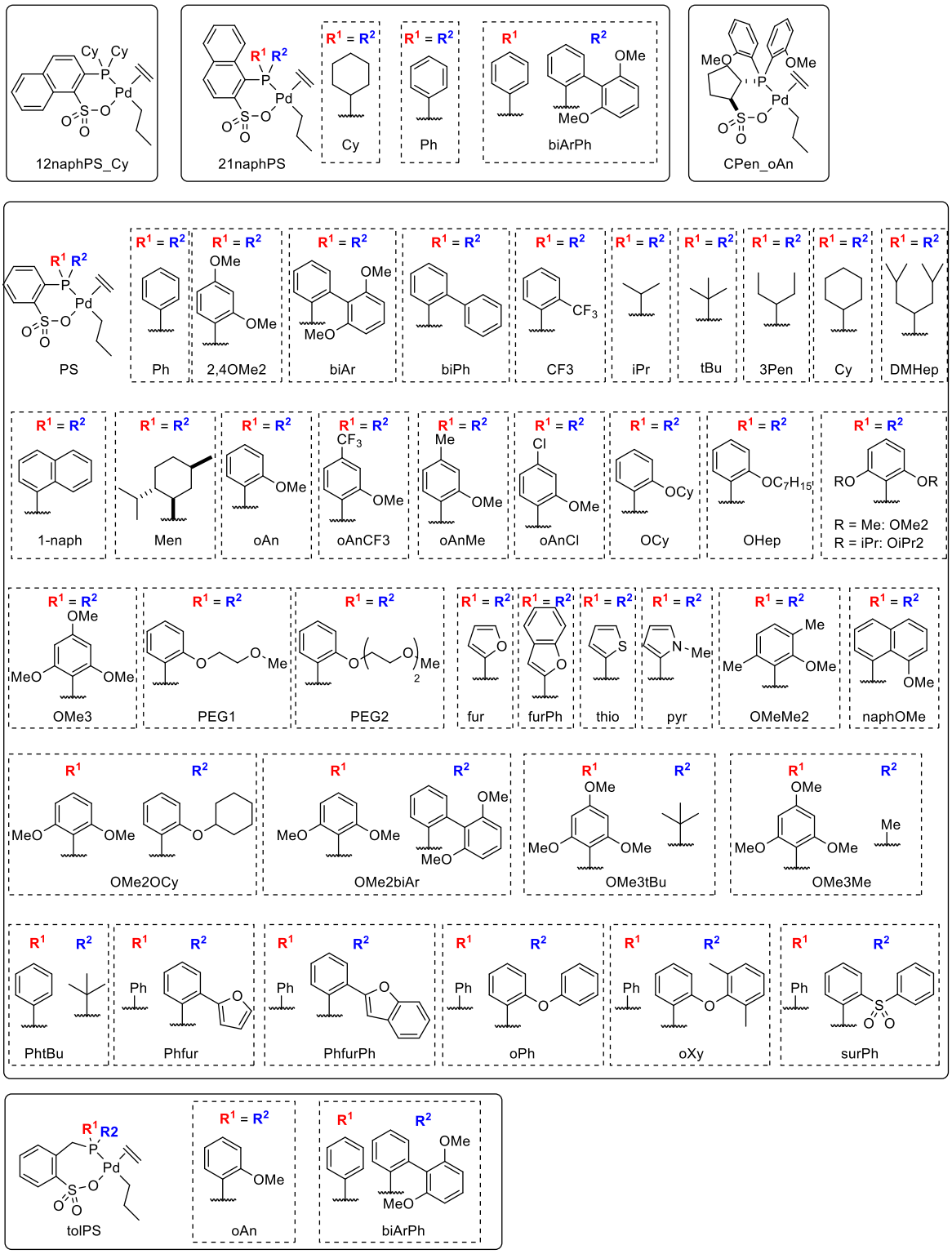


Figure 1. The structures of complexes used in the statistical analysis of polymerizations.

Data and Parameter Collection and Model Construction

Experimental data for ethylene polymerization promoted by methylpalladium complexes bearing PS ligands were collected from the literature. For analysis of both homo- and co-polymerization, the data used for model construction was filtered based on reaction temperature (80 °C for homopolymerization, 80–100 °C for E/MA copolymerization) and split randomly into training/test sets with a ratio of 80:20 (details in supporting information). The same split was used for all analyses. Structures of intermediate **3** with a library of PS ligands (**Figure 1**) were calculated with the geometry optimization method B3LYP-D3(BJ)/[6-31G(d)+LANL2DZ], followed by a single point calculation with B3LYP-D3(BJ)/[6-311+G(d,p)+SDD].

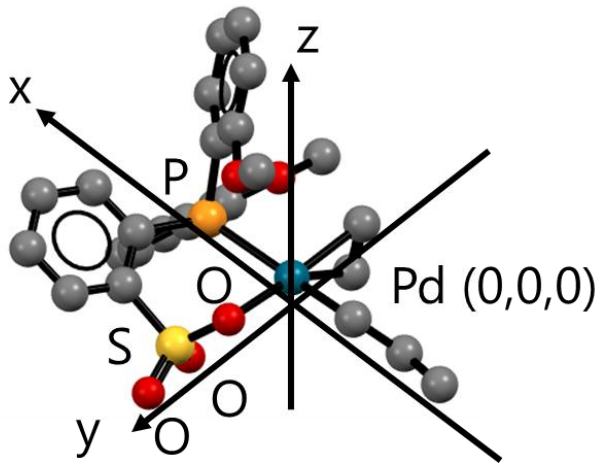


Figure 2. The definition of the coordinate grid.

The CoMFA coordinates were set as described in **Figure 2**. The Pd atom was selected as the origin, and the square planar coordination plane of Pd (i.e., the mean plane of Pd, the two ligating atoms in the PS ligand, and the bound carbon in the propyl group) was defined as the xz plane with the sulfonate group direction as the y

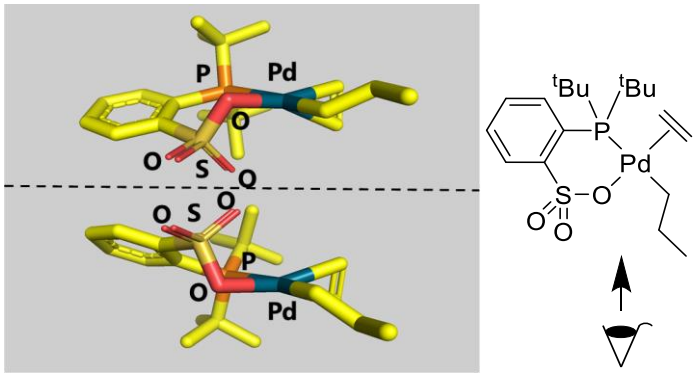


Figure 3. An example of mirror image parallel to coordination plane of palladium.

positive. Although most of the ligands we employed are achiral, the complexes are chiral due to the conformation of the six-membered ring as depicted in **Figure 3**, and as such, both mirror images with respect to the xy plane were equally considered in the analysis. The grid size was set to 1 Å. The information of grid occupation was converted to 10 orthogonal descriptors by principal component analysis (PCA). Conversion of grid occupancy to descriptors and compression by PCA was conducted in Open3dQSAR. Further analysis by linear regression was conducted on R ver. 3.6.1. The results were visualized on the grid with pseudo-coefficients using PyMOL ver. 2.2.3.

A wide variety of descriptors were collected for intermediate **3** from both the simulated complex structures, and their corresponding ligand conformers (**Figure 4**), adding up to a total of 62 parameters (see SI for detail). A few notable examples being structural features (bond lengths, angles, $\%V_{bur}$[12] and Sterimol parameters), orbital information (HOMO and LUMO),[13] natural bond orbital (NBO) analysis of key atoms, the minima of electrostatic potential (V_{min}) localized around the ligating groups, polarizability, chemical shifts, and the stretching frequency for the ethylene bound to palladium.

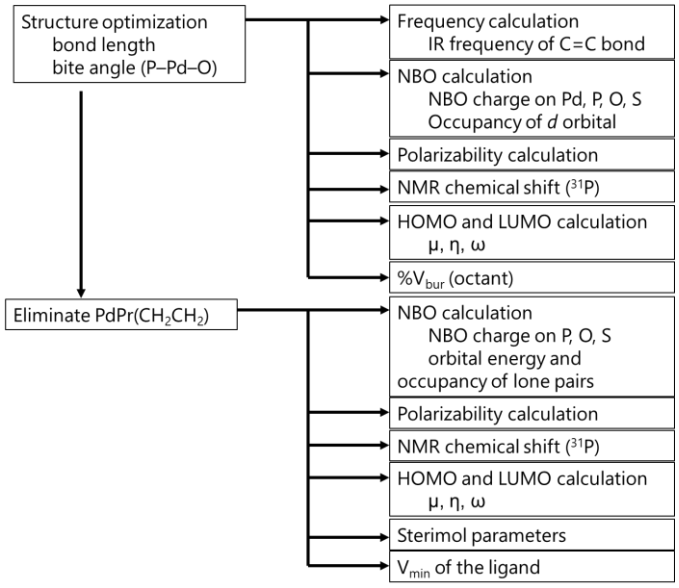


Figure 4. The procedure of parameter collection for machine learning.

The simulated molecular descriptors, along with reaction parameters (e.g., time, temperature, substrate concentration, catalyst loading), were utilized in the analysis of important reaction outputs. For ethylene homopolymerization, statistical models were developed for both molecular weight (MW, the M_n value) and activity, in the forms of log(MW) and log(activity), respectively. For ethylene/methyl acrylate copolymerization, incorporation efficiency (k) was added to the analyzed properties.[14] Incorporation efficiency represents how much the catalyst preferentially incorporates polar monomers during chain propagation, with the assumption that without the bias introduced by the catalyst, incorporation of each monomer is in linear relation to its concentration. A large incorporation efficiency would allow for the direct synthesis of highly functionalized polyolefins with limited concentration of polar monomers during the preparation. Incorporation efficiency k was defined as

$$\text{(incorporation ratio of MA [mol\%])}:\text{(incorporation ratio of ethylene [mol\%])}$$
$$= k\text{(incorporation efficiency)} \times \text{(MA concentration [M])}:\text{(ethylene pressure [MPa]}$$

Thus,

$$k\ \text{(incorporation efficiency)} = \frac{\text{(incorporation ratio of MA [mol\%])} \times \text{(ethylene pressure [MPa])}}{\text{(incorporation ratio of ethylene [mol\%])} \times \text{(MA concentration [M])}} \quad \text{(Eq. 1)}$$

For linear regression, preliminary models were regularized using Akaike's information criterion (AIC). In the random forest analysis, the number of predictors sampled for splitting at each node was tuned by using grid search. Subsequently, the Boruta function was used for elimination of non-explaining descriptors. A 10-fold cross validation was chosen as a validation method. A number of algorithms (e.g., being random forest analysis and support vector machine) were tested for their performances on this dataset (details shown in the Supporting Information).

Comparative Molecular Field Analysis (CoMFA) for Molecular Weight (MW) Analysis

The result of the statistical analysis of log(MW) is shown in **Figure 5a**. The R_{train}^2, the coefficient of determination for train data sets, and R_{test}^2, the coefficient of determination for test data sets, values are 0.89 and 0.78, respectively, suggesting a reasonably robust model. The catalysts in the data set can be divided into two classes: the major class includes ligands that have the same two substituents on the phosphorus atom, while the minor class (9 out of 71 entries) has two different substituents on the phosphorus atom wherein P-atom is a stereogenic center. The model produced a significantly better prediction for the major class and, unsurprisingly, due to limited representation in the training set, the minor class is more poorly predicted.

Partial least squares (PLS) pseudo-coefficient maps can be visualized on the grid. The visualized map on Pd/PS substituted by 2-[2',6'-$(MeO)_2C_6H_3$]C_6H_4 is depicted in **Figure 6**. The red space is positively correlated with an increase in the polymer MW whereas the blue space is the opposite. As marked by a black-dotted circle in the bottom part of the chart, the apical position ~3 Å away from palladium center is highlighted in red, meaning occupation of this position by a substituent is correlative with an increase in the MW of the polymer. The result is consistent with our previous hypothesis.[3a] Namely, the bulkiness of the substituent occupying the apical position is essential to obtaining higher MW, because it blocks the β-hydrogen from the palladium, and thus suppresses the chain-terminating β-hydride elimination. As circled by white-dotted line in **Figure 6**, there is a blue area in the outside of ethylene coordinated to palladium center. This area is occupied when palladium–ethylene bond becomes long thus this should correspond to the weak coordination of ethylene. When the coordination of ethylene is weak in intermediate **3**, ethylene dissociation **TS (3→5)** is accelerated reducing the MW of the polymer.

Aiming to confirm whether extrapolation was possible, data in the limited region was used as training data with the boundary set as log(MW) = 4.25. The result is shown in **Figure 5b**. The trend of the accuracies for each data was similar to the calculation with interpolation shown in **Figure 5a**, while the extrapolation analysis shown in **Figure 5b** contained a relatively larger error. Several data are significantly poorly predicted with examples for the Pd/PS catalyst in which the phosphorus atom is substituted by 2,4,6-$(OMe)_3C_6H_2$ and *t*Bu (The data (exp, pred) = (4.89, 3.83)).

As described above, CoMFA successfully produced a prediction model for the molecular weight of the obtained polymer using a Pd/PS system. It also suggested that 3D information of the catalysts' effect on the MW can be extracted. Next, to explore the effect of catalyst properties more deeply, we probed the system by using DFT-derived properties to build more interpretable statistical models.

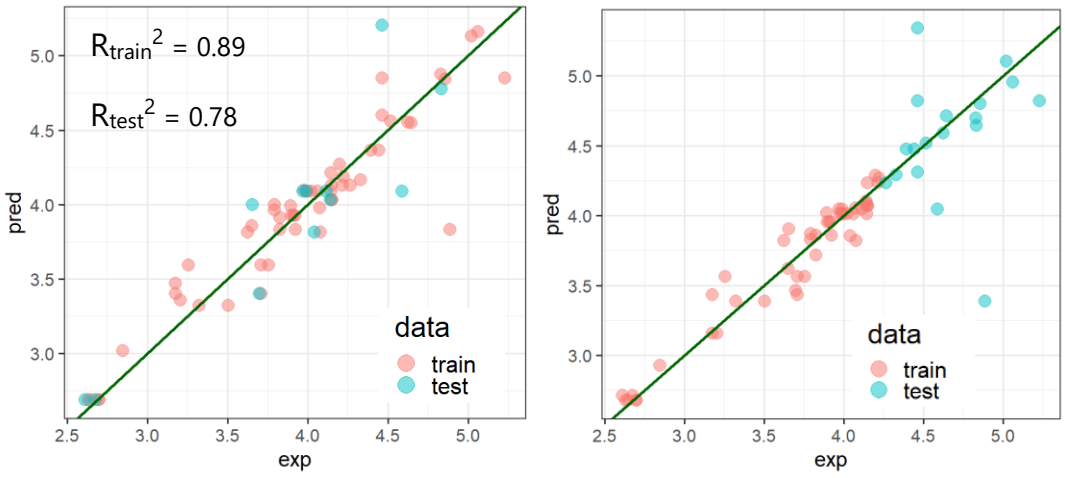


Figure 5. (a) Statistical analysis of molecular weight of ethylene polymerization with CoMFA using PLS and (b) Molecular weight prediction beyond the training set by CoMFA using linear regression.

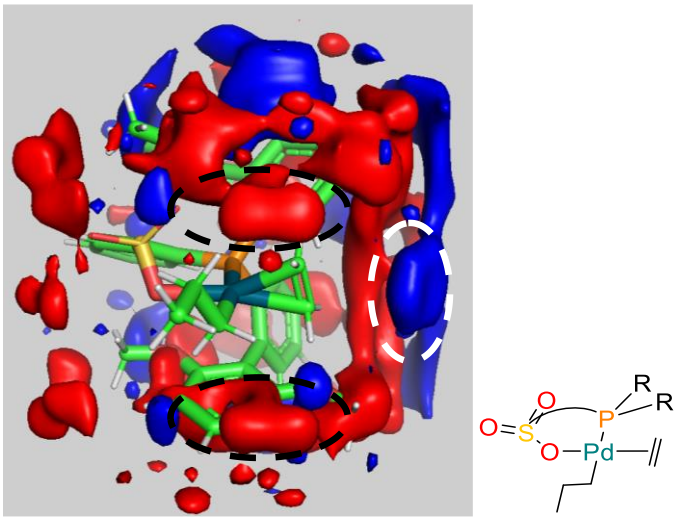


Figure 6. The PLS pseudo-coefficients map drawn on palladium/phosphine–sulfonate complex substituted by 2-[2',6'-$(MeO)_2C_6H_3$]C_6H_4 (red: positive, blue: negative).

Statistical Models for Ethylene Homopolymerization

Using the parameters library as shown in **Figure 4**, we first explored correlating MW for homopolymerization of ethylene. LASSO regression was first used to build a preliminary model (**Figure 7a**), with $R_{train}^2 = 0.84$ and $R_{test}^2 = 0.78$, suggesting function/parameter combination of the model fits for training and test data.

Next, random forest (RF) was chosen for further statistical correlation of the ethylene homopolymerization data set. In this case, not only was the MW of the resulting polymer but also the measured catalytic activity was examined. The calculated models were evaluated via R_{train}^2 and R_{test}^2 from the train and test set, respectively. The resulting statistical analysis of MW of the polymer and polymerization activity using random forest are shown in **Figure 7b** and **Figure 8**, respectively. R_{train}^2 and R_{test}^2 were calculated for each prediction. In both predictions, R_{train}^2 and R_{test}^2 are greater than 0.85 and 0.80 respectively, thus showing good agreement.

The important features impacting the RF fit for the MW analysis are depicted in **Table 1** with the most heavily weighted parameter being the NBO charge on the palladium center. Similarly, the catalytic activity was analyzed. In this case, ^{31}P NMR chemical shifts were found to be the most important parameter (**Table 2**).

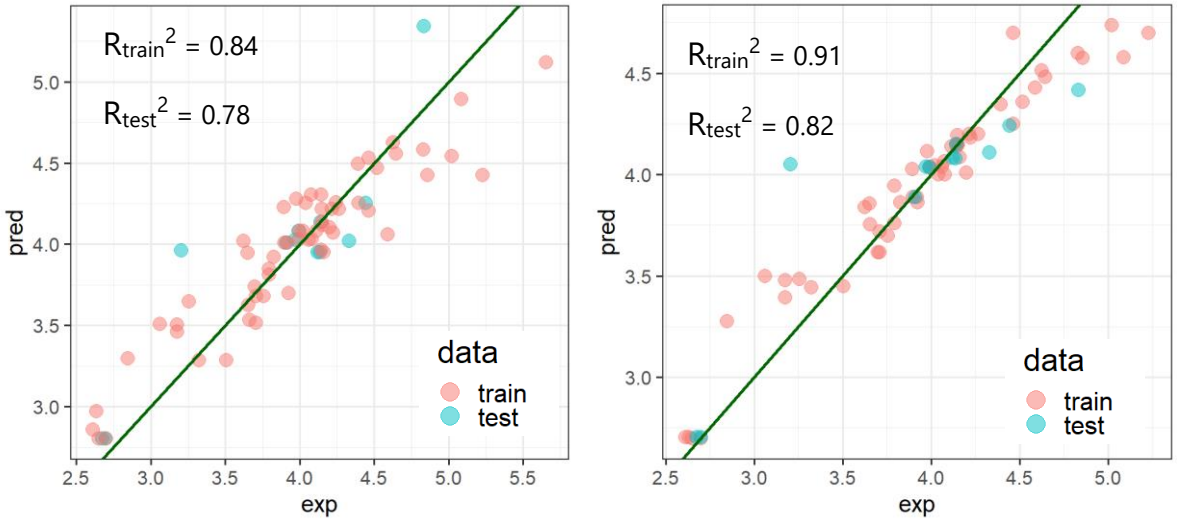


Figure 7. Statistical analysis of log(MW) of ethylene polymerization using (a) LASSO regression and (b) random forest.

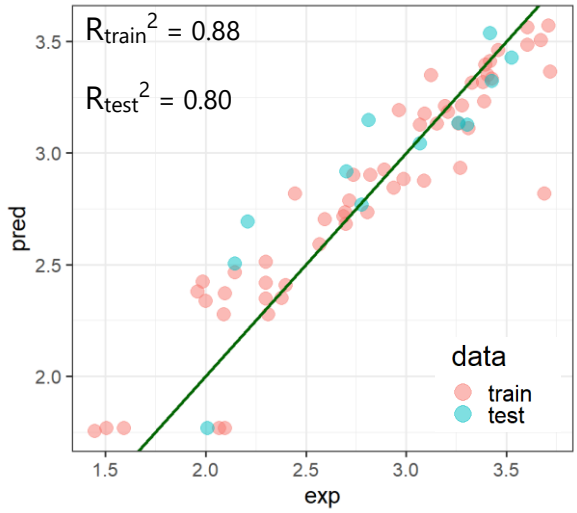


Figure 8. Statistical analysis of activity of ethylene polymerization using random forest.

Table 1. Important parameters for the prediction of the MW of the polymer.

parameter	importance
NBO charge on Pd	1.61
Average of B5 value[a]	1.21
^{31}P NMR chemical shift	1.21
Occupancy of d_{z2}	1.04
NBO charge on S	0.94

[a]average of B5 value of the two substituents on phosphorous atom.

Table 2. Important parameters for the prediction of the polymerization activity of ethylene polymerization.

parameter	importanc
^{31}P NMR chemical shift	1.59
Average of L value[a]	1.50
NBO charge on Pd	1.48
Occupancy of d_{z2}	0.98
Reaction time	0.94

[a]average of L value of the two substituents on phosphorous atom.

Ethylene/Methyl Acrylate Copolymerization

Encouraged by the meaningful analysis on the ethylene homopolymerization, we explored the ethylene/MA copolymerization. First, the MW of the obtained MA/ethylene copolymers was statistically analyzed by LASSO regression (**Figure 9a**). R_{train}^2 and R_{test}^2 were 0.65 and 0.24, respectively, suggesting a poor explanation of the MW in the case of MA/ethylene copolymerization. The random forest approach was further applied to E/MA copolymerization results by Pd/PS. The analyses of three features of the polymerization results, namely the MW of the copolymer, catalytic activities, and incorporation efficiency of MA into the polymer chain are depicted in **Figures 9b**, **c**, and **d**. As shown in **Figures 9b** and **d**, in the cases of MW and MA incorporation efficiency analysis, R_{train}^2 and R_{test}^2 are larger than 0.85 and 0.80 respectively. In contrast, in the case of polymerization activity (**Figure 9c**), R_{test}^2 is lower than 0.75 suggesting the model is insufficient although R_{train}^2 is larger than 0.85. In the incorporation efficiency analysis, R_{train}^2 and R_{test}^2 were 0.91 and 0.82, respectively. For MW of the copolymer, R_{train}^2 and R_{test}^2 were

calculated with randomly separated train/test data sets through 100 iterations. The averages of R_{train}^2 and R_{test}^2 were 0.92 ± 0.009 and 0.73 ± 0.1, suggesting the consistently high validation score supports the validity of the model (See Table S6 for details).

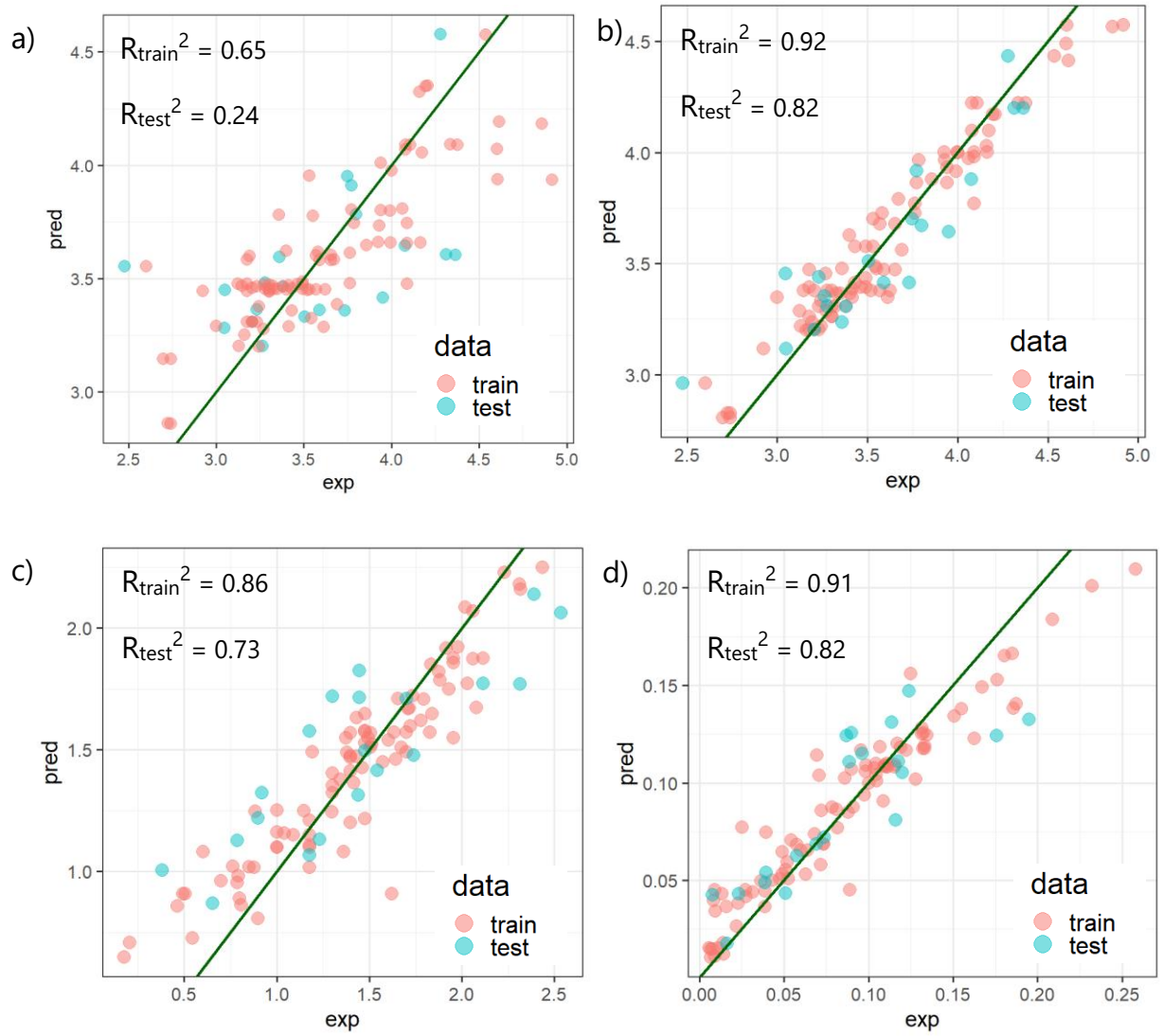


Figure 9. Statistical analysis of MW of E/MA copolymerization using (a) LASSO regression and (b) random forest and statistical analysis of (c) polymerization activity and (d) incorporation efficiency of E/MA copolymerization using random forest.

The important descriptors for MW of the copolymer (**Figure 9b**) are summarized in **Table 3**. The two most important parameters are the average of B5 parameter and $\%V_{bur}$. The important

descriptors for polymerization activity of E/MA copolymerization (**Figure 9c**) are shown in **Table 4**. Finally, incorporation efficiency of MA into the copolymer (defined in Eq. 1, *vide supra*) was analyzed (**Figure 9d**) and the important descriptors are shown in **Table 5**.

Table 3. Important parameters for the prediction of MW of the copolymer

	Importance
Average of B5 parameter	3.20
$\%V_{bur}$	2.02
Occupancy of d_{z^2}	1.92
temperature	1.52
μ^a	1.16

[a]absolute chemical potential

Table 4. Important parameters for the prediction of polymerization activity in E/MA copolymerization.

	Importance
MA concentration	2.82
NBO charge on Pd	1.84
Bite angle	1.57
Ethylene pressure	1.48
Length of Pd–ethylene	1.41

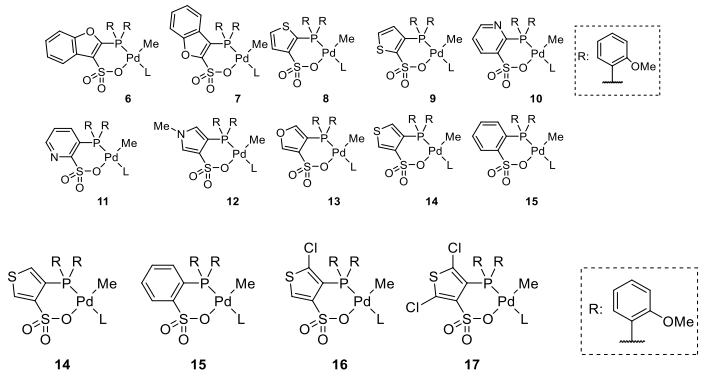
Table 5. Important parameters for the prediction of incorporation efficiency of MA in the copolymer

	Importance
NBO charge on Pd	0.0308
Energy level of lone pair of P	0.0211
MA concentration	0.0210
$\%V_{bur}$	0.0177
Bite angle	0.0172

Prediction of Catalysts Candidates and Synthesis

Based on the above analysis, we designed new Pd/PS-type catalysts as follows. Many of the descriptors describe a trade-off between molecular weight of the polymer and polymerization activity. Also, incorporation of MA decreases the activity. Among the descriptors we examined, we focused our attention on the P–Pd–O bite angle because its importance for polymerization activity was relatively significant but its impact on the polymer molecular weight was predicted to be modest. Accordingly, we assumed that an increase in bite angle would result in higher polymerization activities without a decrease in polymer molecular weight. To increase the bite angle, phosphine-sulfonate complexes bearing five-member ring linkers, **6**–**9** and **12**–**14** were evaluated with our model although it should be noted that this type of catalyst is not included in the training data. Therefore, we were aware that these new catalysts would be considered significant extrapolations of the statistical models. To investigate the effect of the heteroatom, complex **10** and **11** were also considered. The calculated values are shown in Table 6 (vide infra for complexes **16** and **17**). As shown, complex **14** having a 3,4-thenylene linker was predicted to give a polymer molecular weight and polymerization activity comparable to the standard catalyst **15** (Compare Entries 9 and 10). In addition, the incorporation ratio of methyl acrylate with **14** is predicted to be higher than **15**. Based on this prediction, we next synthesized **14**.

Table 6. Predicted and Experimental Results for Palladium Complexes Ligated by Varying Ring Linker Phosphine-sulfonate.[a,b]



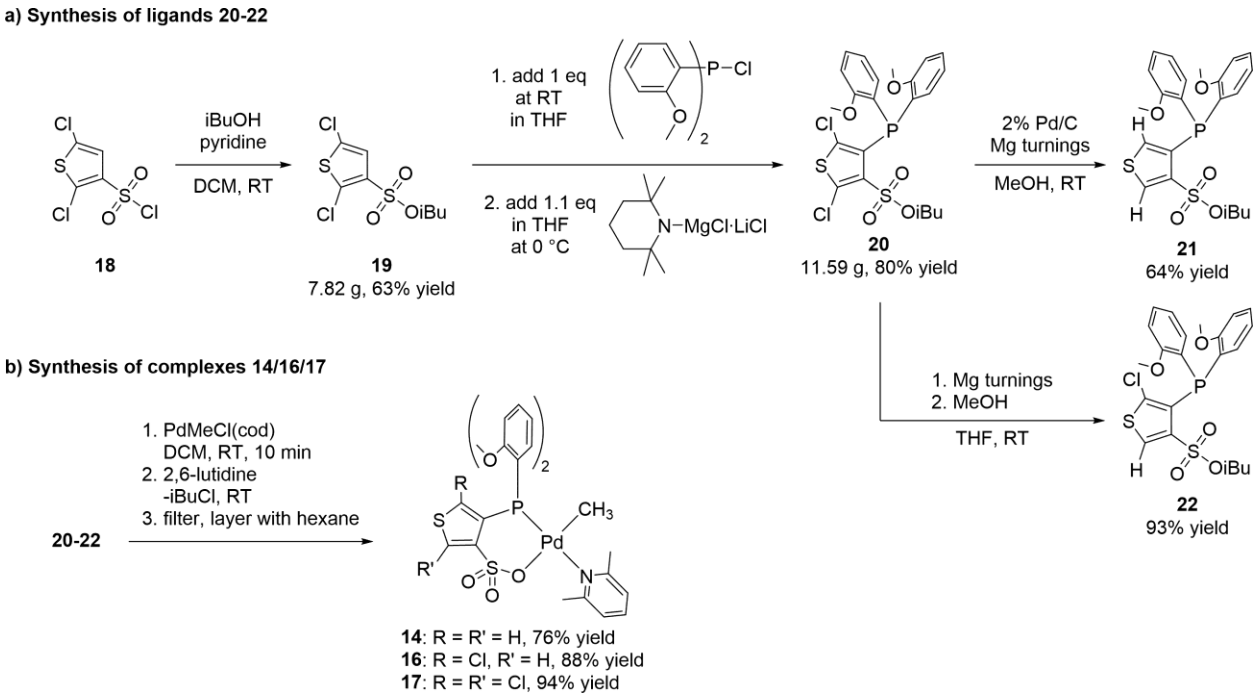
Entry	Cat.	predicted M_n (10^3)	predicted activity (g mmol^{-1} h^{-1})	predicted IR.[c] (mol%)	experimental M_n (10^3)	experimental PDI	experimental activity (g mmol^{-1} h^{-1})	experimental IR.[c] (mol%)
1	**6**	6.1	20	2.5	--	--	--	--
2	**7**	10.3	17	2.1	--	--	--	--
3	**8**	14.8	54	3.6	--	--	--	--
4	**9**	14.0	59	2.0	--	--	--	--
5	**10**	17.4	30	3.9	--	--	--	--
6	**11**	17.4	44	3.8	--	--	--	--
7	**12**	11.1	76	4.5	--	--	--	--
8	**13**	17.6	79	4.3	--	--	--	--
9	**14**	17.6	91	4.4	1.9	2.0	64	3.4
10	**15**	19.0	117	3.4	12.1	2.2	152	5.0
11	**16**	17.6	41.9	2.6	2.5	2.1	339	2.7
12	**17**	17.4	40.3	2.6	6.6	2.1	156	3.8

[a]The prediction was done for the following conditions: Catalyst (10 μmol), methyl acrylate (1.0 M in toluene, total 20 mL) and ethylene (3.0 MPa) in a 50 mL autoclave for 3 h at 80 °C. [b]The polymerizations were done under the following conditions: Catalyst (0.25 μmol in 100 μL CH_2Cl_2), methyl acrylate 3 mL, toluene 11.9 mL and ethylene (3.0 MPa) in a 50 mL autoclave for 1.5 h at 80 °C. [c]IR, incorporation ratio, for prediction was calculated from predicted incorporation efficiency and IR for experiments was determined by quantitative ^{13}C NMR analysis.

Synthesis and structural analyses of 14 and its derivatives 16 and 17

To synthesize complex **14**, the following route towards ligand precursors **20–22** (Scheme 3a) was established: isobutyl ester **19** derived from sulfonyl chloride **18** was mixed with bis(*ortho*-anisyl) chlorophosphine and then treated at 0 °C with TMPMgCl·LiCl, furnishing the dichloro ligand precursor **20**. Double-dehalogenation of **20** using catalytic Pd/C with an excess of magnesium metal in MeOH gave the desired dihydro ligand precursor **21** in 64% isolated yield.[19] Monochloro ligand precursor **22** was also accessed in 93% yield by the selective and quantitative C-Cl magnesiation of the 2-, but not the 5-position of **20**. While the dihydro ligand derivative **21** readily oxidized under air, the electron-deficient chlorinated derivatives **20** and **22** were more oxygen tolerant, which allowed their purification by standard benchtop column chromatography. Ligands **20–22** were first allowed to react with PdMeCl(cod) furnishing intermediate phosphine-sulfonic ester complexes, which upon addition of 2,6-lutidine quantitatively interconverted to the respective palladium complexes **14**/**16**/**17** (Scheme 3b).[19] After crystallization from the crude mixtures, **14/16/17** could be isolated as colorless crystals in 76%, 88% and 94% yield respectively.

Scheme 3. a) Synthesis of ligands **20-22**. b) Synthesis of complexes **14/16/17**.



The molecular structures of complexes **14**, **16** and **17** were further analyzed by single crystal X-ray diffraction (**Figure 10**). Larger P−Pd−O angles were detected in complex **14** (97.08(4)°) as well as monochlorinated **16** (96.30(3)°) when compared to the conventional phenylene linker P/S catalyst **15** (95.00(6)°; CCDC-659797). In contrast, the dichlorinated **17** (86.91(5)°) exhibited a much smaller bite angle. A possible explanation is that, in **17**, the larger chlorine substituent on the 2-position of the thiophene ring causes a rotational restriction of the sulfonate group, whereas for **14** and **16** a smaller hydrogen substituent on the analogous positions allows a rotation and relaxation of the same. The restriction also reflects in the O−S−C3−C2 angles for **14** (54.39(11)°) and **16** (57.46(12)°), which lead to a significantly larger tilt of the Pd-coordinated sulfonate oxygen out of coplanarity with the thiophene unit in comparison with **17** (22.33(19)°). Still all three catalysts show near-perfect square-planar coordination environments with sum-of-angles around the Pd atoms between 359.89° and 360.12° for **14**, **16** and **17**.

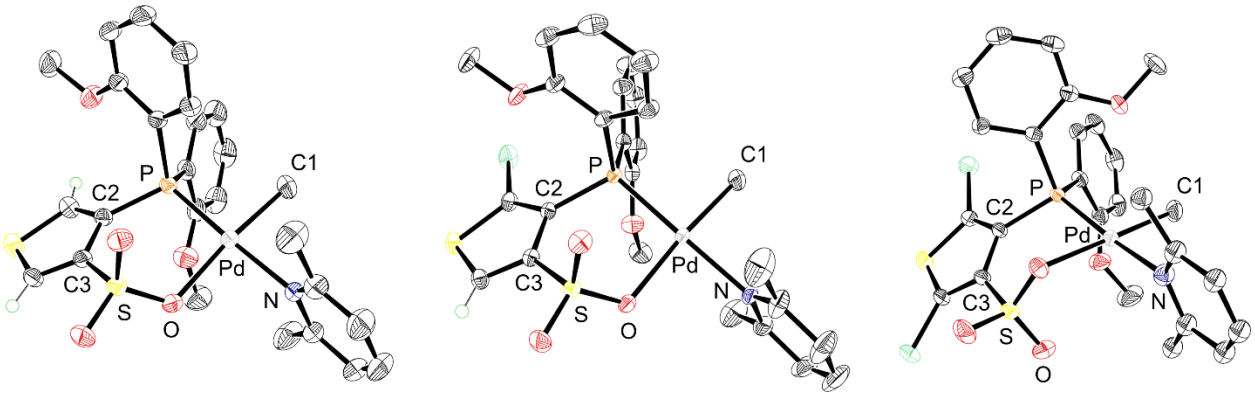


Figure 10. Molecular structure of **14** (left), **16** (middle) and **17** (right) in the crystal (ORTEP drawn at 50% probability; all but the thiophene H atoms and the co-crystalized CH_2Cl_2 molecule in **16** are omitted for clarity; only one disorder position for the 2,6-lutidine ligands in **14/16/17** and a methoxy group in **17** are shown; data for the second disorder position are shown in brackets respectively). Selected bond lengths (Å) and angles (deg) for **14**: Pd−C1 2.026(3), Pd−P 2.2287(5), Pd−O 2.1626(17), Pd−N 2.12(2) (2.131(3)), C1−Pd−P 87.94(6), C1−Pd−N 91.4(9) (89.29(14)), P−Pd−O 97.08(4), O−Pd−N 83.7(9) (85.64(14), O−S−C3−C2 54.39(11), sum-of-angles around Pd (C1/P/O/N): 360.12 (359.95). For **16**: Pd−C1 2.0378(16), Pd−P 2.2468(4), Pd−O 2.1809(10), Pd−N 2.15(3) (2.136(3)), C1−Pd−P 88.66(5), C1−Pd−N 89.5(10) (89.41(11), P−Pd−O 96.30(3), O−Pd−N 85.6(10) (85.62(10)), O−S−C3−C2 57.46(12), sum-of-angles around Pd (C1/P/O/N): 360.06 (359.99). For **17**: Pd−C1 2.028(2), Pd−P 2.2309(6), Pd−O 2.1728(13), Pd−N 2.127(8) (2.121(6)), C1−Pd−P 93.21(8), C1−Pd−N 87.3(2) (87.23(18)), P−Pd−O 86.91(5), O−Pd−N 92.47(19) (92.57(17)), O−S−C3−C2 22.33(19), sum-of-angles around Pd (C1/P/O/N): 359.89 (359.92).

Experimental evaluation of thiophene-based catalysts 14, 16 and 17 and comparison with the predictions

We then applied the new catalysts **14**, **16** and **17** to the copolymerization of methyl acrylate with ethylene and compared their performance with the reference catalyst **15** under the same reaction conditions. The results are summarized in Table 6. The reference catalyst **15** exhibited lower M_n, higher activity (entry 4), and higher incorporation ratio under our experimental conditions (12.1 kg mol^{-1}, 152 kg mol^{-1} h^{-1}, 5.0 mol%), compared to the predicted values in Table 6, Entry 10 (5.2 kg mol^{-1}, 123 kg mol^{-1} h^{-1}, 4.2 mol%). The difference may be attributed to the lower catalyst loading (0.25 μmol) compared to the value used for prediction (10 μmol).

Contrary to our expectations in Table 6 (similar M_n and activity with slightly higher incorporation ratio with **14** compared to **15**), the dihydro complex **14** produced an ethylene/MA copolymer with much lower M_n of 1.9 kg mol^{-1} and catalytic activity of 64 g mmol^{-1} h^{-1} with lower incorporation ratio of 3.4% (Table 6, Entry 9). Of particular note and rather unexpectedly, 5-monochlorinated **16** showed M_n of 2.7 kg mol^{-1} with more than double-fold increase in activity of 339 kg mol^{-1} h^{-1} albeit with a lower MA incorporation of 2.7% (Table 6, compare Entries 9 and 10). The 2,5-di-chlorinated catalyst **17** further increased the M_n to 3.8 kg mol^{-1} with 156 kg mol^{-1} h^{-1} activity and 3.4% incorporation of MA (Table 6, Entry 12). We then calculated the prediction values for **16** and **17** with our model. The results are shown as Entries 11 and 12 in Table 6. The activity for **16** and **17** were both predicted to be lower than **14** or **15**, which is inconsistent with the higher activity experimentally determined using **16**.

Thus, a relationship with the molecular structures from single crystal X-ray crystallography and the catalytic performance as of now in not clear. However, the unexpected observation that increased chlorination of the thiophene linker led to increased catalyst performance for **16** and **17** over **14** suggests why the predictions from the statistical models do not match the observed results. Specifically, the training set mainly contained catalysts with ligand backbones containing 6-membered phenyl linkers, and that P/S catalyst examples containing 5-membered ligand backbones remain rare in literature. Furthermore, the electronic perturbation due to the linker is also not represented in the literature. Therefore, the poor prediction of the experimental performance of the thiophene derived catalysts **14**, **16** and **17** is perhaps not surprising. Nevertheless, the comparable performance of the novel chlorinated thiophene sulfonate catalysts **16** and **17** to the well-established reference system **15** was unexpected and provides a foundation for exploring novel PS type catalysts in the future.

Conclusion

Molecular features impacting the ethylene homopolymerization and copolymerization with MA catalyzed by palladium complexes ligated by phosphine–sulfonate were analyzed by means of a number of data science tools. These analyses led to the identification of ligand features important for the resultant molecular weight of the obtained polymers:1) The larger maximum width of the substituents on the phosphorus atom described by the B5 parameter results in higher molecular weight in both ethylene homopolymerization and ethylene/methyl acrylate copolymerization; 2) the higher electron density on the palladium center causes higher molecular weight by sacrificing catalyst activity in both types of polymerization reactions. While these findings are consistent with the reported tendencies, some new findings were uncovered: 3) The lower occupancy of the palladium d_z2 orbital results in increase of molecular weight and catalyst activity in both types of polymerization reactions. 4) The larger bite angle of the ligand increased activity of ethylene/methyl acrylate copolymerization without affecting the molecular weight. On the basis of these predictions, three Pd/(partially chlorinated) thiophene-based PS complexes were synthesized. Unexpectedly, the chlorinated complexes lead higher catalytic activity in the MA/ethylene copolymerization. One explanation for the experimental derivations from the statistical models likely lies in the under representation of five-membered ring type PS catalysts and halogenated linkes in model training. Lengths/widths of ligands or bite angles are not solely steric factors, but these are hybrid parameters that likely account for electronic impact as well. Future work will exploit this new finding in the development of improved catalysts while also incorporating data to enhance the predictive statistical models.

ASSOCIATED CONTENT

Supporting Information.

The following files are available free of charge.

Detail of the calculations (PDF)

Structural data for DFT calculations (XYZ)

Crystal structures for complexes **14**/**16**/**17** (CIF)

AUTHOR INFORMATION

Corresponding Author

*E-mail for K.N.: nozaki@chembio.t.u-tokyo.ac.jp

*E-mail for M.S.S.: matt.sigman@utah.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

A part of this work was supported by JSPS KAKENHI JP18H05259 and JST ERATO JPMJER2103. The theoretical calculations were performed using computational resources provided by Research Center for Computational Science, National Institutes of Natural Sciences, Okazaki, Japan. S.A. thanks the Program for Leading Graduate Schools (MERIT) from the JSPS. M.S.S. thanks the National Science Foundation (CHE-1763436) for funding.

REFERENCES

(1) Drent, E.; van Dijk, R.; van Ginkel, R.; van Oort, B.; Pugh, R. I. Palladium catalysed copolymerisation of ethene with alkylacrylates: polar comonomer built into the linear polymer chain. *Chem. Commun.* **2002**, 744–745.

(2) For reviews, see: (a) Nakamura, A.; Ito, S.; Nozaki, K. Coordination-Insertion Copolymerization of Fundamental Polar Monomers. *Chem. Rev.* **2009**, *109*, 5215–5244. (b) Ito, S.; Nozaki, K. Coordination-Insertion Copolymerization of Polar Vinyl Monomers by Palladium Catalysts. *Chem. Rec.* **2010**, *10*, 315–325. (c) Nakamura, A.; Anselment, T. M. J.; Claverie, J.; Goodall, B.; Jordan, R. F.; Mecking, S.; Rieger, B.; Sen, A.; van Leeuwen, P. W. N. M.; Nozaki, K. Ortho-Phosphinobenzenesulfonate: A Superb Ligand for Palladium-Catalyzed Coordination-Insertion Copolymerization of Polar Vinyl Monomers. *Acc. Chem. Res.* **2013**, *46*, 1438–1449. (d) Carrow, B. P.; Nozaki, K. Transition-Metal-Catalyzed Functional Polyolefin Synthesis: Effecting Control through Chelating Ancillary Ligand Design and Mechanistic Insights. *Macromolecules* **2014**, *47*, 2541–2555. (e) Baier, M. C.; Zuideveld, M. A.; Mecking, S. Post-Metallocenes in the Industrial Production of Polyolefins. *Angew. Chem. Int. Ed.* **2014**, *53*, 9722–9744. (f) Ito, S. Chain-Growth Polymerization Enabling Formation/Introduction of Arylene Groups into Polymer Main Chains. *Polym. J.* **2016**, *48*, 667–677. (g) Ito, S. Palladium-Catalyzed Homo- and Copolymerization of Polar Monomers: Synthesis of Aliphatic and Aromatic Polymers. *Bull. Chem. Soc. Jpn.* **2018**, *91*, 251–261. (h) Tan, C.; Chen, C. Emerging Palladiumand Nickel Catalysts for Copolymerization of Olefins with Polar Monomers. *Angew. Chem. Int. Ed.* **2019**, *58*, 7192–7200.

(3) For representative reports since 2014, see: (a) Ota, Y.; Ito, S.; Kuroda, J.; Okumura, Y.; Nozaki, K. Quantification of the Steric Influence of Alkylphosphine–Sulfonate Ligands on Polymerization,

Leading to High-Molecular-Weight Copolymers of Ethylene and Polar Monomers. *J. Am. Chem. Soc.* **2014**, *136*, 11898–11901. (b) Jian, Z.; Wucher, P.; Mecking, S. Heterocycle-Substituted Phosphinesulfonato Palladium(II) Complexes for Insertion Copolymerization of Methyl Acrylate. *Organometallics* **2014**, *33*, 2879–2888. (c) Contrella, N. D.; Jordan, R. F. Lewis acid modification and ethylene oligomerization behavior of palladium catalysts that contain a phosphine-sulfonate-diethyl phosphonate ancillary ligand. *Organometallics* **2014**, *33*, 7199–7208. (d) Jian, Z.; Mecking, S. Insertion Homo- and Copolymerization of Diallyl Ether. *Angew. Chem. Int. Ed.* **2015**, *54*, 15845–15849. (e) Chen, M.; Yang, B.; Chen, C. Redox-Controlled Olefin (Co)Polymerization Catalyzed by Ferrocene-Bridged Phosphine-Sulfonate Palladium Complexes. *Angew. Chem. Int. Ed.* **2015**, *54*, 15520–15524. (f) Schuster, N.; Rünzi, T.; Mecking, S. Reactivity of Functionalized Vinyl Monomers in Insertion Copolymerization. *Macromolecules* **2016**, *49*, 1172–1179. (g) Jian, Z.; Leicht, H.; Mecking, S. Direct Synthesis of Imidazolium-Functional Polyethylene by Insertion Copolymerization. *Macromol. Rapid Commun.* **2016**, *37*, 934–938. (h) Wu, Z.; Chen, M.; Chen, C. Ethylene Polymerization and Copolymerization by Palladium and Nickel Catalysts Containing Naphthalene-Bridged Phosphine–Sulfonate Ligands. *Organometallics* **2016**, *35*, 1472–1479. (i) Jian, Z.; Mecking, S. Insertion Polymerization of Divinyl Formal. *Macromolecules* **2016**, *49*, 4395–4403. (j) Jian, Z.; Falivene, L.; Boffa, G.; Sánchez, S. O.; Caporaso, L.; Grassi, A.; Mecking, S. Direct Synthesis of Telechelic Polyethylene by Selective Insertion Polymerization. *Angew. Chem. Int. Ed.* **2016**, *55*, 14378–14383. (k) Wei, J.; Shen, Z.; Filatov, A. S.; Liu, Q.; Jordan, R. F. Self-Assembled Cage Structures and Ethylene Polymerization Behavior of Palladium Alkyl Complexes That Contain Phosphine-Bis(arenesulfonate) Ligands. *Organometallics* **2016**, *35*, 3557–3568. (l) Ota, Y.; Ito, S.; Kobayashi, M.; Kitade, S.; Sakata, K.; Tayano, T.; Nozaki, K. Crystalline Isotactic Polar Polypropylene from the Palladium-Catalyzed Copolymerization of

Propylene and Polar Monomers. *Angew. Chem. Int. Ed.* **2016**, *55*, 7505–7509. (m) Wada, S.; Jordan, R. F. Olefin Insertion into a Pd-F Bond: Catalyst Reactivation Following β-F Elimination in Ethylene/Vinyl Fluoride Copolymerization. *Angew. Chem. Int. Ed.* **2017**, *56*, 1820–1824. (n) Chen, M.; Chen, C. Rational Design of High-Performance Phosphine Sulfonate Nickel Catalysts for Ethylene Polymerization and Copolymerization with Polar Monomers. *ACS Catal.* **2017**, *7*, 1308–1312. (o) Yang, B.; Pang, W.; Chen, M. Redox Control in Olefin Polymerization Catalysis by Phosphine–Sulfonate Palladium and Nickel Complexes. *Eur. J. Inorg. Chem.* **2017**, 2510–2514. (p) Liang, T.; Chen, C. Side-Arm Control in Phosphine-Sulfonate Palladium- and Nickel-Catalyzed Ethylene Polymerization and Copolymerization. *Organometallics* **2017**, *36*, 2338–2344. (q) Wu, Z.; Hong, C.; Du, H.; Pang, W.; Chen, C. Influence of Ligand Backbone Structure and Connectivity on the Properties of Phosphine-Sulfonate Pd(II)/Ni(II) Catalysts. *Polymers* **2017**, *9*, 168. (r) Black, R. E.; Jordan, R. F. Synthesis and Reactivity of Palladium(II) Alkyl Complexes that Contain Phosphine-cyclopentanesulfonate Ligands. *Organometallics* **2017**, *36*, 3415–3428. (s) Yang, B.; Xiong, S.; Chen, C. Manipulation of polymer branching density in phosphine-sulfonate palladium and nickel catalyzed ethylene polymerization. *Polym. Chem.* **2017**, *8*, 6272–6276. (t) Zhang, D.; Chen, C. Influence of Polyethylene Glycol Unit on Palladium- and Nickel-Catalyzed Ethylene Polymerization and Copolymerization. *Angew. Chem. Int. Ed.* **2017**, *56*, 14672–14676. (u) Song, G.; Pang, W.; Li, W.; Chen, M.; Chen, C. Phosphine-sulfonate-based nickel catalysts: ethylene polymerization and copolymerization with polar-functionalized norbornenes. *Polym. Chem.* **2017**, *8*, 7400–7405. (v) Black, R. E.; Jordan, R. F. Synthesis and Reactivity of Palladium(II) Alkyl Complexes That Contain Phosphine-Cyclopentanesulfonate Ligands. *Organometallics* **2017**, *36*, 3415–3428. (w) Zou, C.; Pang, W.; Chen, C. Influence of

Chelate Ring Size on the Properties of Phosphine-Sulfonate Palladium Catalysts. *Sci. China Chem.* **2018**, *61*, 1175–1178.

(4) Wucher, P.; Goldbach, V.; Mecking, S. Electronic Influences in Phosphinesulfonato Palladium(II) Polymerization Catalysts. *Organometallics* **2013**, *32*, 4516–4522.

(5) (a) Verloop, A. In *Drug Design*; Ariens, E. J., Ed.; Academic Press: New York, 1976; Vol. 3, p 133. (b) Verloop, A. In *QSAR in Drug Design and Toxicology*; Hadzi, B. J.-B., Ed.; Elsevier: Amsterdam, 1987; p 97. (c) 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. (d) Ehm, C.; vittoria, A.; Gorynov, G. P.; Izmer, V. V.; Kononovich, D. S.; Samsonov, O. V.; Di Girolamo, R.; Budzelaar, P. H. M.; Voskoboynikov, A. Z.; Busico, V.; Uborsky, D. V.; Cipullo, R. An Integrated High Throughput Experimentation/Predictive QSAR Modeling Approach to ansa-Zirconocene Catalysts for Isotactic Polypropylene. *Polymers*, **2020**, *12*, 1005.

(6) The relationship between log(MW) and ΔΔG can be written as

$$\Delta G(\text{TS for chain transfer}) - \Delta G(\text{TS for ethylene insertion}) = RT\ln(DP)$$

$$= \frac{RT}{\log e}\log(\text{MW}) - RT\ln(28)$$

(DP: degree of polymerization, (MW of polymer)/(MW of monomer). This means $\Delta\Delta G[\text{kcal/mol}] = 1.6 \times \log(\text{MW}) - 2.3$ at 80 °C.

(7) Hostaš, J.; Řezáč, J. Accurate DFT-D3 Calculations in a Small Basis Set. *J. Chem. Theory Comput.* **2017**, *13*, 3575–3585.

(8) (a) 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*, 6424. (b) Reid, J. P.; Sigman, M. S. Holistic Prediction of Enantioselectivity in Asymmetric Catalysis. *Nature* **2019**, *571*, 343–348.

(9) Nakano, R.; Chung, L. W.; Watanabe, Y.; Okuno, Y.; Okumura, Y.; Ito, S.; Morokuma, K.; Nozaki, K. Elucidating the Key Role of Phosphine−Sulfonate Ligands in Palladium-Catalyzed Ethylene Polymerization: Effect of Ligand Structure on the Molecular Weight and Linearity of Polyethylene. *ACS Catal.* **2016**, 6101–6113.

(10) Santiago, C. B.; Guo, J.-Y.; Sigman, M. S. Predictive and mechanistic multivariate linear regression models for reaction development. *Chem. Sci.* **2018**, *9*, 2398–2412.

(11) Martin, Y. C. Quantitative Drug Design; Marcel Dekker: New York, 1978.

(12) Clavier, H.; Nolan, S. P. Percent buried volume for phosphine and *N*-heterocyclic carbeneligands: steric properties in organometallic chemistry. *Chem. Commun.* **2010**, *46*, 841–861.

(13) The HOMO and LUMO orbital images are shown in supporting information.

(14) The case of incorporation ratio << 1 was reported: Nakano, R.; Nozaki, K. Copolymerization of Propylene and Polar Monomers Using Pd/IzQO Catalysts. *J. Am. Chem. Soc.* **2015**, *137*, 10934–10937.