Catalysis Science & Technology

PAPER

View Article Online
View Journal | View Issue

Cite this: *Catal. Sci. Technol.*, 2014, **4**, 979

A theoretical study of the activity in Rh-catalysed hydroformylation: the origin of the enhanced activity of the π-acceptor phosphinine ligand†

Sonia Aguado-Ullate,[a] John A. Baker,[b] Vanessa González-González,[a] Christian Müller,[c] Jonathan D. Hirst*[b] and Jorge J. Carbó*[a]

Received 22nd November 2013,
Accepted 6th January 2014

DOI: 10.1039/c3cy00956d

www.rsc.org/catalysis

The factors governing the activity in Rh-catalyzed hydroformylation were investigated using a set of computational tools. We performed DFT calculations on the phosphinine-modified Rh catalyst [$HRh(CO)_3(PC_5H_2R_3)$] and compared it to the phosphane-modified $HRh(CO)_3(PR_3)$ and $HRh(CO)_2(PR_3)_2$ complexes. The π-acceptor phosphinine ligand coordinates preferentially at the equatorial site of the pentacoordinated Rh complex with the heterocycle perpendicular to the equatorial plane, although the ligand freely rotates around the Rh–P bond. The overall energy barrier can be divided into the following contributions: alkene complex formation, alkene rotation and alkene insertion. In the absence of steric effects (model systems), the overall barrier correlates with the computed barrier for alkene rotation. This proves that π-acceptor ligands reduce back-donation to the alkene, leading to a lower rotational barrier and, consequently, to a higher activity. The Rh–P donor–acceptor interactions were quantified using a modified version of energy decomposition analysis (EDA). In Rh–phosphinine systems, the efficient directionality of the π-back-donation, rather than the overall acceptor ability, is responsible for the high catalytic activity. Introducing steric effects increases the energy required to coordinate the alkene, increasing the overall barrier. The factors governing the activity in Rh–monophosphane catalysts seem to be related to those derived for Rh–diphosphane during the development of a QSAR model (*Catal. Sci. Technol.* 2012, **2**, 1694). To investigate whether the findings for mono- can be extrapolated to diphosphane ligands, we re-examined our previous QSAR model using the Topological Maximum Cross Correlation (TMACC) method based on easy-to-interpret 2D-descriptors. The TMACC descriptors highlight heteroatoms close to phosphorus as activity-increasing atoms, whereas highly substituted carbon atom groups are highlighted as activity-decreasing groups.

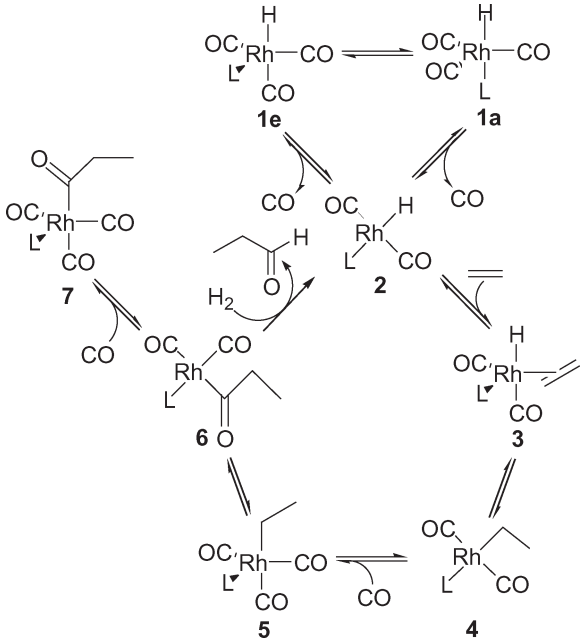
Introduction

The hydroformylation of alkenes is one of the largest applications of homogeneous catalysis in industry.[1] It consists of the addition of carbon monoxide and hydrogen to alkenes, catalyzed by cobalt, rhodium or platinum catalysts, leading to the formation of aldehydes. For industrial production, the phosphane-modified rhodium catalysts are widely used because they can show high activity and regioselectivity towards the usually desired linear aldehyde.[1,2]

Scheme 1 shows the generally accepted mechanism for the hydroformylation of alkenes catalyzed by phosphane-modified rhodium catalysts.[3] Different reaction kinetics have been observed experimentally depending on ligand properties that, in general, fit into one of two extreme cases. For electron-poor ligands such as bulky monophosphites and the unmodified rhodium–carbonyl catalysts, the hydrogenolysis of the acyl species **6** controls the overall rate of hydroformylation.[4,5] In contrast, the kinetics for electron-rich ligands are consistent with a rate-determining step early in the catalytic cycle. Recently, a combination of isotope effects study and computational analysis have demonstrated that the overall process from the resting state species **1** to hydride migration (from complex **3** to complex **4**) governs the overall activity in 1-octene hydroformylation catalyzed by the rhodium–xantphos complexes.[6] Independently, a theoretical study by Jensen and co-workers led to the same conclusion for phosphane and moderately

[a] *Departament de Química Física i Inorgànica, Universitat Rovira i Virgili, Campus Sescelades, C/Marcel.lí Domingo s/n, 43007 Tarragona, Spain. E-mail: j.carbo@urv.cat*
[b] *School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, UK. E-mail: jonathan.hirst@nottingham.ac.uk*
[c] *Institut für Chemie und Biochemie, Freie Universität Berlin, Fabeckstr. 34/36, 14195 Berlin, Germany*

† Electronic supplementary information (ESI) available: Tables containing the relative energy values for coordination isomers of mono- and bi-phosphane complexes, the detailed results and description of the EDA bond analysis, and the *xyz* coordinates for the most relevant computed structures. See DOI: 10.1039/c3cy00956d



Scheme 1

electron-withdrawing phosphite ligands whereas for strongly electron-withdrawing ligands, calculations supported hydrogenolysis as the rate-determining step.[7]

Recent clarification of the rate determining step[6,7] indicates that the ligands that promote CO dissociation, alkene coordination, or hydride migration might yield higher catalytic turnovers. Besides this, some systematic studies have attempted to establish correlations between the ligand structure and the catalytic activity.[8–11] Early experimental studies on monodentate P-donor ligands showed that a relationship exists between ligand basicity and the catalytic activity; thus, the least basic phosphanes enhance the activity.[8] In addition, van Leeuwen and co-workers showed that phenoxophosphane (phosphacyclic) moieties are less basic than diphenylphosphino moieties and exhibit an increase in the catalytic rate.[9] In general, for the ligands enclosing a phosphorus atom inside a cycle, an increase of the activity was observed and was attributed to a lower basicity.[12]

The differences in the catalytic activity have also been related to the steric features of the ligands. For a series of phosphinine ligands, the variation in the catalytic performance was attributed to the steric properties of the ligand.[13] According to the authors, the bulky phosphinine ligands favour the formation of a monoligand rhodium species, which should have a larger accessible space compared with a diligand rhodium species. Similarly, under industrial conditions, an excess of the phosphane ligand is used because the selectivity towards a linear aldehyde is improved; however, the activity is reduced because the phosphane dissociation equilibrium shifts from the monophosphane $[HRh(CO)_3(PR_3)]$ towards the less active $HRh(CO)(PPh_3)_2$ species. On the other hand, for a series of diphosphane xantphos-type ligands,[14] the rate increases with increasing bite angle,[15] while we showed that increasing the bite angle increases the steric hindrance around the metal center.[16] During the development of the quantitative structure–activity relationship (QSAR) model for the hydroformylation catalyzed by Rh–diphosphane complexes, we discovered that complex relationships underlie the origin of the activity and that both the shape and the electronic properties of the catalyst need to be considered.[17]

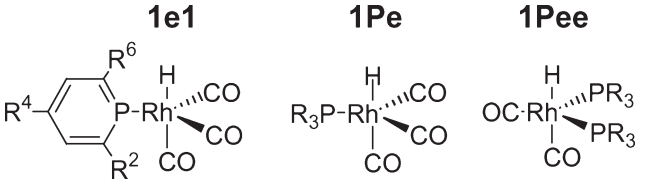
Here, we focus on the study of the phosphinine-modified rhodium catalyst $[HRh(CO)_3(PC_5H_2R_3)]$, for which computational investigations are still lacking,[18] and compared it with the phosphane-modified catalysts $[HRh(CO)_3(PR_3)]$ and $[HRh(CO)_2(PR_3)_2]$ (Scheme 2). The Rh–phosphinine system showed a much higher activity compared to classical PPh_3-based catalysts[13,19] and followed an analogous kinetics.[13] Our aim is to understand at the molecular level the factors governing the activity, evaluating them and their interplay. In addition, we investigate whether the findings for monophosphane ligands can be extrapolated to diphosphane, re-examining a previous QSAR model[17] by using the easy-to-interpret TMACC descriptors.[20]

Results and discussion

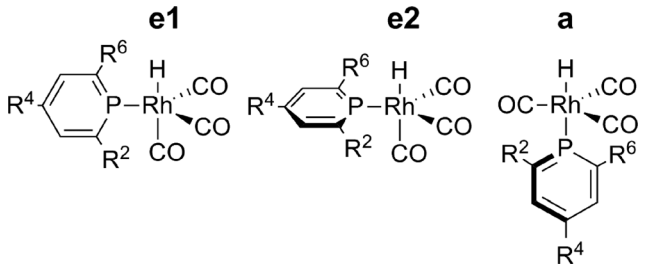
Coordination preferences of the phosphinine ligand

The coordination and interaction with rhodium of phosphane ligands have been extensively studied by computational methods,[21,22] but to the best of our knowledge there are no such studies on phosphinine ligands. Thus, initially, we analyzed the coordination mode of $[HRh(CO)_3(2,4,6\text{-}PC_5H_2R_3)]$ (R = H, Ph) complexes, which corresponds to the resting states in Rh-catalyzed hydroformylation. The phosphinine ligand can coordinate in the equatorial (e) and apical (a) positions. It is also possible to generate two additional isomers for the e coordination depending on whether the heterocycle is perpendicular to the equatorial plane of the complex or if it is in-plane (**1e1** and **1e2**, respectively, in Scheme 3 and Fig. 1).

Previous high pressure NMR studies on the analogous and more stable iridium(I) system indicated that only one phosphinine ligand is coordinated to the transition metal center under hydroformylation conditions.[13] These experiments also suggested an equatorial position for the phosphinine, presumably as two rotamers. The DFT calculated relative energies support the experimental proposal. For the $[HRh(CO)_3(PC_5H_5)]$ model complex, the most stable isomer is $\mathbf{1e1^H}$ followed by $\mathbf{1e2^H}$ and $\mathbf{1a^H}$ (+1.1 and +1.4 kcal mol^{-1} higher in energy, respectively). The preference for equatorial coordination with the phosphinine cycle perpendicular to



Scheme 2



Scheme 3

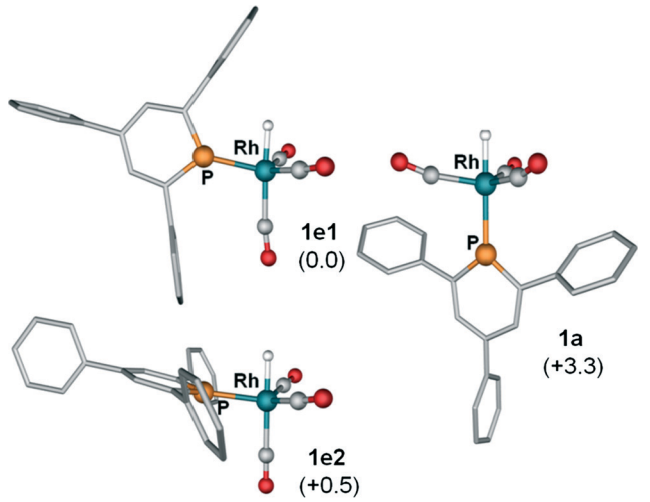


Fig. 1 Calculated 3D molecular structures of the different geometric and conformational isomers **1e1**, **1e2** and **1a** of the $[HRh(CO)_3(PC_5H_2Ph_3)]$ complex. Relative energies are in kcal mol^{-1}.

the equatorial plane of the complex can be explained by electronic arguments based on frontier molecular orbitals (FMO). The in-plane d_{xy} orbital of the metal fragment is high in energy and hybridized away from the other equatorial ligands favouring π-back-donation to the phosphorus p orbital perpendicular to the ligand plane.[23] This is reflected in the shortest Rh–P distance for $\mathbf{1e1^H}$ (2.37 Å *vs.* 2.39 and 2.41 Å for $\mathbf{1e2^H}$ and $\mathbf{1a^H}$, respectively).

Introducing the ligand steric effects *via* calculations on $[HRh(CO)_3(2,4,6\text{-}PC_5H_2Ph_3)]$ complexes did not change the order of the relative energies (0.0, +0.5 and +3.3 kcal mol^{-1} for **1e1**, **1e2** and **1a**, respectively). Nevertheless, the energy difference between the two equatorial rotamers diminishes, while between equatorial and apical coordination it increases. Both trends can be attributed to the steric effects. The ideal P–Rh–CO angle in **1e1** and **1a** is 90°, whereas in **1e2** it is 120° (see Fig. 1). Thus, the first two isomers, **1e1** and **1a**, with a smaller angle are slightly destabilized with respect to **1e2** by the steric interactions between the phosphinine substituents and the auxiliary ligands.

In order to analyze the rotation around the Rh–P bond in e compounds, we performed a relaxed energy scan of the H–Rh–P–C_{ortho} dihedral angle. The calculations did not show any barrier connecting the rotational isomers $\mathbf{1e1^H}$ and $\mathbf{1e2^H}$. This, along with the small energy difference between the rotamers, even for the 2,4,6-triphenylphosphinine system, strongly indicates that the Rh–P bond rotates freely. This might hamper the application of axially chiral monodentate phosphinines in the asymmetric hydroformylation of prochiral substrates.[24] As a matter of fact, the preferred equatorial coordination would place the chiral centers far away from the apical region, in which the key ligand–substrate interactions should take place.[17,25,26] For example, ligands with axial chirality such as binaphos, which shows excellent performance in Rh-catalyzed asymmetric hydroformylation,[27] induce enantioselectivity *via* interactions between the substrate and the axially chiral groups of the apical ligand moiety.[26] Moreover, the very low barrier computed for ligand rotation would allow easy ligand reorganization upon interaction with the incoming substrate, leading to low enantiodiscrimination.

Next, we analyzed the coordination preferences of the phosphinine ligand in the rate-determining step, in which the transition state for alkene insertion into the Rh–H bond is involved (Fig. 2). Using ethene as a model substrate, the computed equatorial path for the model system, $\mathbf{TSe1^H}$, is lower in energy than the apical path, $\mathbf{TSa^H}$, by 2.0 kcal mol^{-1}, increasing somewhat the energy difference found in the resting-state (1.4 kcal mol^{-1}). For equatorial coordination, the attempts to locate a transition state with the phosphinine parallel to the equatorial plane ended in the corresponding TS for path **e1**. The perpendicular disposition of the ligand favours back-donation from Rh and, in turn, reduces back-donation to the alkene, which then can easily rotate to reach the TS for insertion. If we recall the low rotational barrier for the Rh–P bond, it is reasonable to think that in the case that the alkene complex **3e2** is formed, the system would tend to switch to the lower-energy easily-accessible **e1** reaction channel. Thus, calculations indicate that most of the reaction will occur through the channel with the phosphinine in an equatorial position and perpendicular to the equatorial plane.

Analysis of the overall energy barrier for the Rh–phosphinine and phosphane systems

The overall energy barrier can be computed as the energy difference between the transition state for alkene insertion (TS)

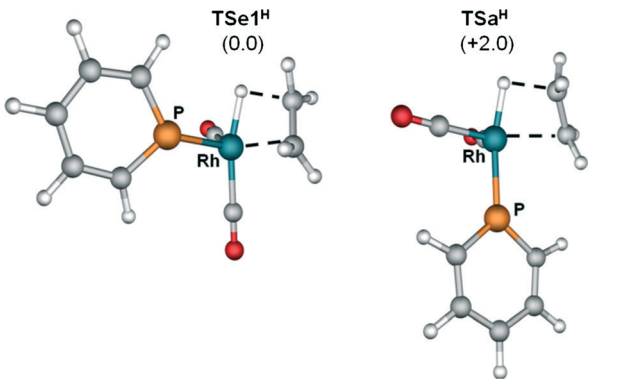


Fig. 2 Molecular structures of the transition states for alkene insertion on phosphinine–Rh systems: $\mathbf{TSe1^H}$ and $\mathbf{TSa^H}$. Relative energies are in kcal mol^{-1}.

and the rhodium hydride carbonyl resting-state of catalyst **1**.[6,7] To identify the individual factors governing the activity, we decomposed the overall barrier into the energetic cost of several steps: the formation of the alkene complex (ΔE_{alkene}) and the energy barriers for alkene rotation ($\Delta E^{\neq}_{rot}$) and insertion ($\Delta E^{\neq}_{inser}$). Note that the alkene insertion proceeds through the rotation of the alkene moiety out of the equatorial plane of the trigonal bipyramidal (bpt) complex **3**, followed by the transfer of the hydride moiety to generate the alkyl species. Table 1 collects the values of the most favourable path for each system, and Fig. 3 shows some of the key structures. For mono- and bicoordinated PPh_3 systems the lowest energy paths are the equatorial and the equatorial–equatorial paths (see the ESI†). Interestingly, at the ONIOM level,[22f] in which the electronic effects of Ph substituents are neglected, the equatorial–apical path is more favoured. This can be rationalized as follows: assuming similar steric interactions at both levels, the electronic effects of the Ph groups reduce the donation ability of the phosphane and, consequently, its tendency to be placed in the apical position.[23,26] Thus, although the results need to be viewed cautiously, it is more straightforward to compare the equatorial paths.

The overall energy barriers reproduce the experimental observations,[13,19] providing a clear picture of the activity differences. The phosphinine system shows a lower value (19.0 kcal mol^{-1}) compared to the monophosphane complex (22.4 kcal mol^{-1}), while the bis(phosphane)-based complex shows the highest value (25.9 kcal mol^{-1}) (see Table 1). For monocoordinated phosphinine and phosphane complexes, the overall barrier follows the same trend as the alkene insertion and rotation barrier, with the energy increasing by ~3 kcal mol^{-1} on going from phosphinine to phosphane. This indicates that the alkene rotation process governs the overall barrier, and it is responsible for the observed higher activity of the Rh–phosphinine system. Furthermore, it proves previous statements suggesting that for electron-withdrawing ligands the amount of back-donation is small, leading to facile rotation of the alkene moiety and therefore a low barrier for alkene insertion.[7,21]

Following the previous arguments, for electron-donor bis(phosphane) systems one would expect higher rotational barriers and, consequently, higher overall barriers. This is observed for the model system, in which both barriers increase by ~1 kcal mol^{-1} with respect to the monophosphane system (Table 1, values in parentheses). On the other hand,

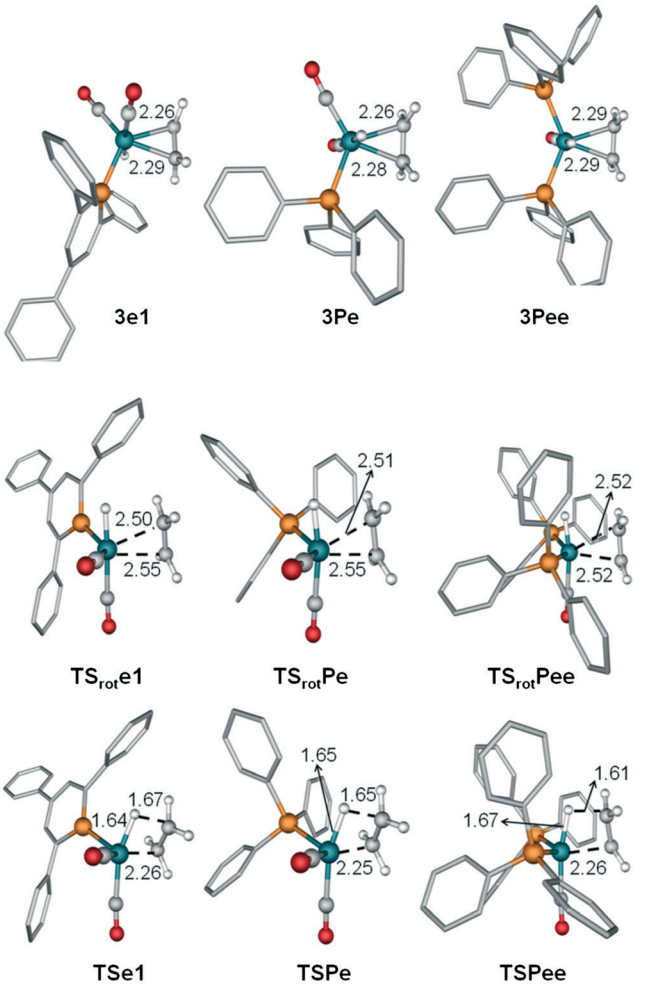


Fig. 3 Molecular structures of the alkene complexes $[HRh(CO)_n(L)_n(H_2CCH_2)]$ **3e1**, **3Pe** and **3Pee** and the transition states for alkene rotation (**TS$_{rot}$**) and insertion into the Rh–H bond (**TS**). Ligand hydrogen atoms are omitted for clarity. Distances are in Å.

for bulky PPh_3 phosphanes the alkene rotation barrier decreases upon bi-coordination, whereas the overall barrier is still high. Introducing the bulky groups increases the steric repulsion between the alkene substrate and the equatorial PR_3 ligands in **3Pee** (see Fig. 3). This was reflected in longer Rh–alkene carbon distances for **3Pee** than for **3Pe** (2.287 *vs.* 2.272 Å on average) and in smaller alkene interaction energies (−24 and −30 kcal mol^{-1}); therefore, the more loosely bound alkene in **3Pee** can rotate more easily. On the other hand, the formation of the alkene complex **3Pee** is energetically more costly than the formation of **3Pe** (15.1 *vs.* 11.5 kcal mol^{-1}). The latter effect dominates, explaining the higher overall barrier. Thus, the increase of the overall barrier upon coordination of the second phosphane is not a direct consequence of the electronic properties of the ligands but of their steric properties. This is in line with our previous findings which indicated that the activity of Rh-catalyzed hydroformylation depends on both the basicity and the shape of the ligand[17] and with the proposed rate-controlling factors by Jensen and co-workers.[7] Thus, we can propose for the design of new active

Table 1 Overall ($\Delta E^{\neq}_{overall}$), alkene insertion ($\Delta E^{\neq}_{inser}$) and rotation ($\Delta E^{\neq}_{rot}$) energy barriers and alkene coordination energy ($\Delta E^{\neq}_{alkene}$)[a]

	Phosphinine	Mono-phosphane	Bi-phosphane
$\Delta E^{\neq}_{overall}$ (**1** → **TS**)	19.0 (19.3)	22.4 (21.9)	25.9 (22.5)
$\Delta E^{\neq}_{inser}$ (**3** → **TS**)	7.7 (10.8)	10.9 (13.5)	10.8 (12.3)
$\Delta E^{\neq}_{rot}$ (**3** → **TS$_{rot}$**)	4.5 (5.9)	7.1 (7.3)	5.6 (8.6)
$\Delta E^{\neq}_{alkene}$ (**1** → **3**)	11.3 (8.5)	11.5 (8.5)	15.1 (10.2)

[a] Energies are in kcal mol^{-1} for the $[HRh(CO)_3(PC_5H_2R_3)]$, $[HRh(CO)_3(PR_3)]$ and $[HRh(CO)_2(PR_3)_2]$ systems (R = Ph; the values in parentheses are for R = H).

ligands that reducing their steric hindrance and increasing their π-acidity will favour alkene coordination and rotation to reach the insertion TS, causing a reduction of the overall energy barrier and higher catalytic activities.

Rhodium–phosphorus and –alkene bond analysis: the origin of the enhanced activity of π-acceptor ligands

To evaluate the electronic properties of the P-ligands, we can use the IR stretching frequencies of the CO ligand (ν_{CO}) in the *trans*-$L_2Rh(CO)Cl$ complex. They follow the order of $P(OPh)_3 > P(OMe)_3 >$ 2,4,6-triphenylphosphinine $> PPh_3 > PEt_3$.[28] A large ν_{CO} value indicates π-acceptor properties due to reduced π-back-donation from the metal center to the CO ligand, while a small value is indicative of strong σ-donation. Thus, these values show that phosphinines are poorer electron donors than phosphanes but richer than phosphites. Besides the overall electronic donating ability of the ligand, it is interesting to consider the decomposition into the individual contributions, σ-donation and π-back-donation, as was done in the classical Dewar–Chatt–Duncanson model. To evaluate these contributions we used a modified version of energy decomposition analysis (EDA) based on an orbital deletion procedure, which allows the bonding to be broken into physically meaningful components (see the ESI† and Fig. S1 for details).[29] Fig. 4 schematically describes the orbital interactions in the analysis of the Rh–P bonding. We focused on the model systems because steric effects are put aside.

Table 2 collects the main results of the EDA bonding analysis for the Rh–P and –alkene ligands and the full analysis is shown in Table S2 of the ESI.† We also show the values for the Rh–phosphite bond in the $[Rh(CO)_2H(PH_3)_{ap}(P(OH)_3)_{eq}]$ complex[30] obtained at the same computational level as in a previous study.[26] For the P-ligands we observed that the interaction energies (ΔE_{int}) follow the same trend of the CO stretching frequencies: $P(OH)_3 > PC_5H_5 > PH_3$ (−29.6, −24.8 and −20.8 kcal mol^{-1}, respectively). The values of σ donation [ΔE_σ(L → Rh)] are very similar for all the ligands (~13 kcal mol^{-1}). On the other hand, the π back-donation energy decreases within the series (−10.3, −8.5, and −6.9 kcal mol^{-1} for $P(OH)_3$, PC_5H_5 and PH_3, respectively), explaining the higher overall charge donation from the ligand to the metal. The analysis of the Rh–alkene bond nicely correlates with rotational barriers and with the donor/acceptor properties on going from phosphinine to mono-phosphane and to bis(phosphane). Thus, this analysis quantifies the electronic factors and proves that the π-acceptor property enhances the activity in Rh-catalyzed hydroformylation.

Interestingly, for phosphinine π-back-donation occurs preferentially in the equatorial plane of the complex to the out-of-plane p_π-type orbitals of phosphorus ($\Delta E_\pi(a'')$ in Table 2). This directionality would lead to a more effective competition for metal electron density with the alkene, resulting in activities close to those of electron-poorer ligands such as phosphites. Accordingly, the computed overall barriers for the $HRh(CO)_3(PC_5H_5)$ and $HRh(CO)_3(P(OH)_3)$ systems are similar, 19.3 and 20.6 kcal mol^{-1}, respectively. Thus, although the overall π-acceptor property in phosphinines is somewhat reduced with respect to phosphites,[13] the efficient directionality of π-back-donation in phosphinines leads to very active catalysts.

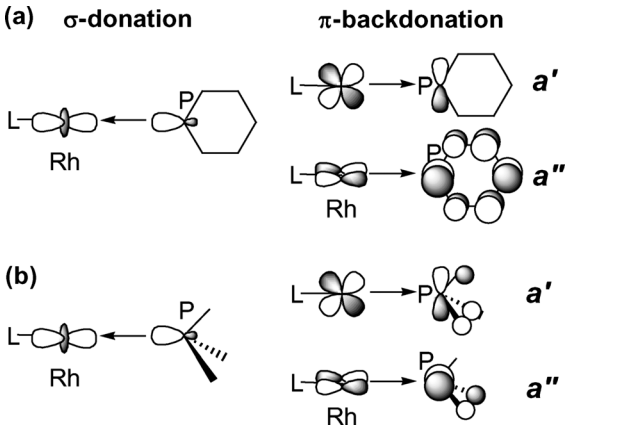


Fig. 4 Representation of the interactions between the transition metal fragment and the phosphinine (a) and phosphane (b) ligands in terms of the Dewar–Chatt–Duncanson model.

Table 2 Interaction energies (ΔE_{int}) and σ-donation/π-backdonation evaluation for the Rh–phosphorus and Rh–alkene bonds[a]

	Rh–phosphorus			Rh–alkene		
	$P(OH)_3$–	PC_5H_5 **1e1^H**	PH_3 **1Pe^H**	**3e1^H**	**3Pe^H**	**3Pee^H**
ΔE_{int}	−29.6	−24.8	−20.8	−30.4	−33.9	−35.9
ΔE_σ(L → Rh)	−13.1	−13.5	−13.6	−14.2	−13.7	−13.8
ΔE_π(Rh → L)	−10.3	−8.5	−6.9	−22.6	−24.7	−27.1
$\Delta E_\pi(a')$	—	−3.2	−3.2	—	—	—
$\Delta E_\pi(a'')$	—	−5.3	−3.7	—	—	—

[a] Energies are in kcal mol^{-1}. For $P(OH)_3$, the values are taken from ref. 26 for the $[HRh(CO)_2(PH_3)(P(OH)_3)]$ complex with the $P(OH)_3$ ligand in the equatorial position. For **1e1^H** and **1Pe^H**, the calculations imposed the C_s symmetry to decompose the π-backdonation into in-plane (a') and out-of-plane (a'').

Factors governing the catalytic activity: correlation with QSAR models for diphosphane ligands

In this section, we analyzed whether the findings for mono-phosphanes can be extrapolated to diphosphane ligands. As stated in the introduction, we have discussed the factors governing the activity for Rh–diphosphane catalysts during the development of a 3D-QSAR model.[17] We have observed the correlation between high activity and low basicity for a given subset of structurally related complexes; however, when comparing ligands of similar basicity the shape of the catalyst seems to determine the activity differences.[17] Nevertheless, the use of alignment-independent 3D-descriptors made the chemical interpretation of the mathematical model difficult.[31] Thus, here we re-examined the previous QSAR model using the Topological Maximum Cross Correlation (TMACC) method based on easy-to-interpret alignment-independent 2D-descriptors.[20] During recent years, QSAR approaches have emerged as an alternative in the theoretical study of catalysis,[32] including those based on alignment-independent descriptors.[17,33]

The dataset defined from the experimental work in ref. 34 consists of 19 diphosphane ligands (set 1: **L1–L19**, see Fig. 5) and uses percentage of conversion (%conv.) as the response variable. Unfortunately, we could not expand the model including monophosphane ligands because there are no suitable data available that allow one to compare the conversion values. We considered four atomic properties: Gasteiger partial charges[35] to represent electrostatics, Crippen–Wildman molar refractivity[36] to represent steric properties and, in addition, Crippen–Wildman log P parameters[36] representing lipophilicity, and the log S parameter[37] representing solubility. Table 3 collects the statistical parameters of the 10-fold cross-validation. All of the defined individual descriptors yielded models that are close to the limit of the prediction ability ($0.55 < q^2 < 0.65$), a model with $q^2 < 0.5$ being considered non-predictive. When we combined the four properties, an acceptable model was obtained with $q^2 = 0.71$. A closer inspection of the data revealed that ligand **L13** has the largest difference between the experimental and predicted conversion values, 61%; consequently, it could be classified as an outlier. The analysis of the chemical space showed that all ligands with intermediate or high activities have heteroatoms in their structures except for **L13**. This means that the features of this type of structure are probably not well represented in the training data. When we set aside ligand **L13** (set 2), the statistical parameters improved significantly ($q^2 = 0.89$ and $r^2 = 0.92$). These findings indicate that the structure–activity relationship requires sophisticated descriptors that include electronic and steric factors.

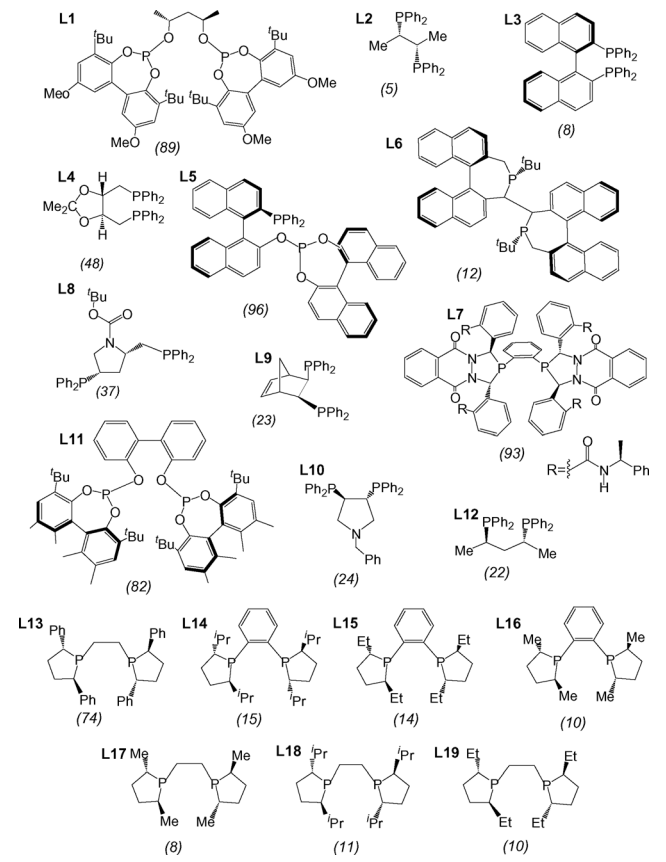


Fig. 5 The ligand dataset and activity outcome (%conv.) for hydroformylation of styrene by rhodium complexes.

Table 3 Statistical parameters of the 10-fold cross-validation for the activity using different types of descriptors[a]

	Set 1 (19 ligands)		Set 2 (18 ligands)	
Descriptor	q^2	r^2	q^2	r^2
Electrostatic	0.59	0.64	0.72	0.78
Steric	0.64	0.69	0.72	0.76
Lipophilicity	0.58	0.78	0.76	0.91
Solubility	0.59	0.89	0.78	0.99
Combination of all	0.71	0.76	0.89	0.92

[a] The Pearson correlation coefficient (r^2) and the cross-validated coefficient of determination (q^2).

Besides, the TMACC descriptors provide a method for interpreting the results when combined with a linear regression method such as PLS. The predicted activity of the ligand can be partitioned among its constituent atoms. Fig. 6 shows the representation of the interpretation of the diphosphite ligands **L1** and **L11**, the phosphane–phosphite **L5**, and the diazophospholane **L7** that have the highest activities (%conv. > 80) in the dataset. The TMACC method allows the visualization of the atoms that contribute most positively to the activity in blue and those that decrease the activity the most in red (yellow and orange colors represent intermediate positive and negative contributions). The oxygen atoms of the phosphite moieties and the hydrazine groups of the diazophospholane ligand are coloured blue (Fig. 6), and hence have been identified as activity-increasing groups. On the other hand, disubstituted sp^2 carbons involving terminal alkyl groups and some other highly substituted carbons are colored in red (Fig. 6), indicating activity-decreasing groups. These substituted carbons may be related with ligand bulkiness, which, in turn, is related with the reduction of the catalyst activity. Fig. 7 displays the interpretation of the electrostatic and the steric property

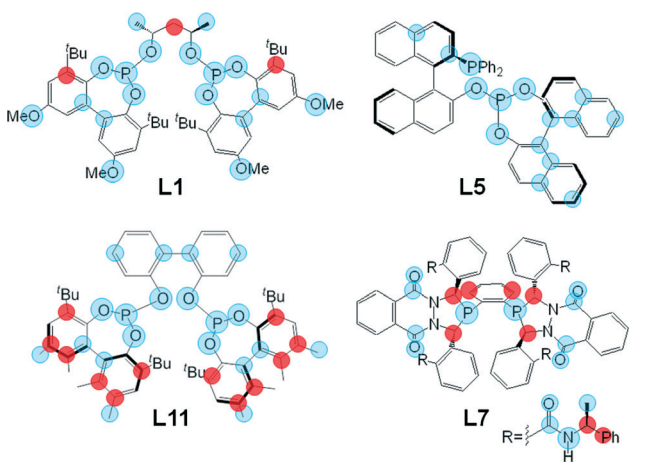


Fig. 6 Visualization of the TMACC/PLS QSAR models for set 2 combining all the descriptors derived for ligands **L1** (chiraphite), **L5** (binaphos), **L7** (diazophospholane) and **L11** (kelliphite). In the interest of clarity, we did not show the lower contribution of other atoms. High positive contributions to activity are shown in blue and high negative contributions are shown in red.

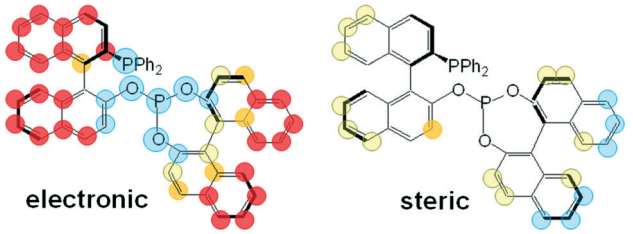


Fig. 7 Visualization of the TMACC/PLS QSAR models for set 2 built from electrostatic and steric properties for catalytic **L5** (binaphos). High and intermediate positive contributions to activity are shown in blue and yellow, and high and intermediate negative contributions are shown in red and orange.

models for ligand **L5**. The change in coloring for these properties indicates that electrostatics has a more dramatic effect on the activity of the ligand: the heteroatoms change from blue in the electrostatic-based model to uncolored in the steric-based model. A similar pattern was observed for ligands **L1**, **L7** and **L11**, indicating that for them the high activity is also dominated by the electrostatic properties induced by the heteroatoms. To sum up, the interpretation of the QSAR model for diphosphane ligands can be related to our findings for monophosphane, indicating that similar rules govern the activity of both types of ligands.

Conclusions

We analyzed the behaviour of the π-acceptor phosphinine ligand in rhodium-catalyzed hydroformylation of alkenes and compared it to phosphane-modified catalysts $[HRh(CO)_3(PR_3)]$ and $[HRh(CO)_2(PR_3)_2]$. The $2,4,6\text{-}PC_5H_2Ph_3$ phosphinine ligand coordinates preferentially at the equatorial site of the pentacoordinated rhodium complex with the heterocycle perpendicular to the equatorial plane of the complex, although the ligand freely rotates around the Rh–P bond.

We divided the overall energy barrier into several steps and/or contributions (alkene complex formation, alkene rotation and alkene insertion), which were then evaluated. In the absence of steric effects (model systems), the overall barrier correlates with the barrier for alkene rotation. This proves that for π-acceptor ligands the amount of back-donation to the alkene is small, leading to its facile rotation and, consequently, to a higher activity. We also quantified the donor/acceptor interactions of the Rh–P bonds using a modified version of EDA analysis. Although the overall π-acceptor ability of phosphinines is lower than that of other ligands such as phosphites, the efficient directionality of their in-plane π-back-donation leads to very active catalysts. Introducing the steric effects of the ligands causes an increase of the energy required to form the alkene complex and, consequently, an increase of the overall barrier. The factors governing the activity in Rh–monophosphane catalysts are closely related to those of Rh–diphosphane catalysts. This was confirmed by re-examining a previous QSAR model using the easy-to-interpret TMACC descriptors.

Thus, the design of active ligands in rhodium-catalyzed alkene hydroformylation should increase their π-acidity and reduce their steric hindrance. Nevertheless, one should be aware that bulky monodentate ligands could favor the formation of $[HRh(CO)_3L]$ complexes which are more active than the $[HRh(CO)_2L_2]$ ones.

Experimental section

Computational details: DFT calculations

The DFT calculations were carried out using the Amsterdam density functional program (ADFv2008).[38] The electronic configurations were described by a triple-ζ plus polarization Slater-type basis set, as included in the ADF package. The 1s–3d electrons for Rh, the 1s electrons for C and O, and the 1s–2p electrons for P were treated as frozen cores. We applied scalar relativistic corrections to them *via* the zeroth-order regular approximation (ZORA) with the core potentials generated using the DIRAC program.[38] We used the GGA functional BP86.[39,40] Full geometry optimization without any symmetry constraints was performed, unless otherwise stated. The transition states were characterized by a single imaginary frequency and the normal mode, which corresponds to the expected reaction path. In some structures, we found a residual imaginary frequency related with the loose torsion angles of the bulky substituents. We are aware that most of the popular DFT methods such as BP86, B3LYP or PBEh are unable to describe noncovalent interactions in their attractive regime.[41] In a recent study of Rh-catalyzed hydroformylation,[26] we tested the M06 class of functionals recommended to study noncovalent interactions[42] and the B97D functional including explicit dispersion corrections.[43] Both functional types gave qualitatively the same results as the BP86 one for these systems in which the ligand–substrate interactions were dominated by repulsive interactions.[26] We expected the same results here. Indeed, the interactions between the phenylphosphino moieties and aliphatic alkenes were proved to be repulsive in nature for alkene insertion into the Rh–H bond by means of QM/MM calculations.[16]

To analyze the nature of the Rh–phosphorus and –alkene bonds, we employed the Energy Decomposition Analysis (EDA) method[44] and a modified version based on orbital deletion that allows separating the σ and the π interactions in a physically meaningful manner.[29] The details are provided in the ESI.†

QSAR TMACC-based modelling

These descriptors are generated using atomic properties (electrostatics, solubility, steric effect and lipophilicity) determined by molecular topology. The source code for computing TMACC descriptors is available for download at http://comp.chem.nottingham.ac.uk/download/TMACC. The electrostatic properties are represented by the *Gasteiger partial charge*,[35] which is calculated using the method of partial equalization of orbital electronegativity. This procedure

calculates the atomic charges in σ-bonded and non-conjugated π-systems using only the topology of the catalysts. The *Crippen–Wildman molar refractivity* (MR) is used as a measure of the steric effects that is determined through the classification of atoms based on the neighbouring atoms.[36] The *Crippen–Wildman partition coefficients* ($\log P$) are assigned to each atom as a measure of atomic lipophilicity, determined in the same way as the *Crippen–Wildman molar refractivity* (MR).[36] The solubility properties are described by $\log S$ parameters, representing solubility and solvation phenomena.[37] We scaled each contribution by the largest absolute value, so that the positive and negative values took maximum values of +1 and −1.

The TMACC autocorrelation descriptor (x_{ac}) is given as

$$x_{ac}(p,d) = \sum p_i p_j$$

where p is one of the properties and d is the topological distance between atoms i and j, normally the shortest number of bonds between the atoms. The sum is over all atom pairs that are separated by a distance d. We treat each atomic property that can take positive and negative values as separate properties. The molar refractivity is the exception that only takes positive values. Like for the GRIND descriptor, we keep only the maximum value calculated for any given distance. For the purposes of interpretation (see above), for each descriptor, we recorded which atoms contributed to the maximum product. In the event of more than one pair having the same value, we record all pairs. The maximum distance was as large as the largest distance in the molecule. The minimum distance was zero; that is, we allowed i = j.

We employed Partial Least Squares (PLS) Regression[45] as the multivariate regression technique. Ten-fold cross validation was used for model building and evaluation. Different statistical parameters facilitated the evaluation of the predictive ability of the models during the fitting and test stages, namely the Pearson correlation coefficient (r^2) and the cross-validated correlation coefficient (q^2) (see the ESI†).[46] The TMACC descriptor string for each molecule comprises several hundred elements. Clearly, it is not practicable to present the resultant QSAR equation, in which the latent variables in the PLS regression are linear combinations of the descriptors. A more qualitative summary of the model is provided *via* the graphical interpretations in Fig. 6 and 7, in which the PLS has effectively been reversed to ascribe contributions to the activity to individual atoms.

Acknowledgements

The authors are grateful for financial support from the MICINN of Spain (CTQ2011-29054-C02-01), from the Direcció General de Recerca (DGR) of the Autonomous Government of Catalonia (grants 2009SGR462 and XRQTC) and from EPSRC (EP/I006559).

Notes and references

1 R. Franke, D. Selent and A. Börner, *Chem. Rev.*, 2012, **112**, 5675.
2 *Rhodium Catalyzed Hydroformylation*, ed. P. W. N. M. van Leeuwen and C. Claver, Kluwer Academic Publishers, Dordrecht, Netherlands, 2000.
3 D. Evans, J. A. Osborn and G. Wilkinson, *J. Chem. Soc. A*, 1968, 3133.
4 (*a*) M. Garland and P. Pino, *Organometallics*, 1991, **10**, 1693; (*b*) V. S. Nair, S. P. Mathew and R. V. Chaudhari, *J. Mol. Catal. A: Chem.*, 1999, **143**, 99; (*c*) J. Feng and M. Garland, *Organometallics*, 1999, **18**, 417.
5 P. C. J. Kamer, A. van Rooy, G. C. Schoemaker and P. W. N. M. van Leeuwen, *Coord. Chem. Rev.*, 2004, **248**, 2409.
6 E. Zuidema, L. Escorihuela, T. Eichelsheim, J. J. Carbó, C. Bo, P. C. J. Kamer and P. W. N. M. van Leeuwen, *Chem.–Eur. J.*, 2008, **14**, 1843.
7 M. Sparta, K. J. Børve and V. R. Jensen, *J. Am. Chem. Soc.*, 2007, **129**, 8487.
8 (*a*) P. W. N. M. van Leeuwen and C. F. Roobeek, *J. Organomet. Chem.*, 1983, **258**, 343; (*b*) W. R. Moser, C. J. Papite, D. A. Brannon and R. A. Duwell, *J. Mol. Catal.*, 1987, **41**, 271.
9 E. Zuidema, P. E. Goudriaan, B. H. G. Swennenhuis, P. C. J. Kamer, P. W. N. M. van Leeuwen, M. Lutz and A. L. Spek, *Organometallics*, 2010, **29**, 1210.
10 M. F. Haddow, A. J. Middleton, A. G. Orpen, P. G. Pringle and R. Papp, *Dalton Trans.*, 2009, 202.
11 O. Diebolt, H. Tricas, Z. Freixa and P. W. N. M. van Leeuwen, *ACS Catal.*, 2013, **3**, 128.
12 (*a*) K. Nozaki, N. Sakai, T. Nanno, T. Higashijima, S. Mano, T. Horiuchi and H. Takaya, *J. Am. Chem. Soc.*, 1997, **119**, 4413; (*b*) S. Breeden, D. J. Cole-Hamilton, D. F. Foster, G. J. Schwarz and M. Wills, *Angew. Chem., Int. Ed.*, 2000, **39**, 4106; (*c*) T. P. Clark, C. R. Landis, S. L. Freed, J. Klosin and K. A. Abboud, *J. Am. Chem. Soc.*, 2005, **127**, 5040; (*d*) T. Hayashi, M. Tanaka and I. Ogata, *J. Mol. Catal.*, 1979, **6**, 1; (*e*) C. F. Hobbes and W. S. Knowles, *J. Org. Chem.*, 1981, **46**, 4422.
13 B. Breit, R. Winde, T. Mackewitz, R. Paciello and K. Harms, *Chem.–Eur. J.*, 2001, **7**, 3106.
14 L. A. van der Veen, P. H. Keeven, G. C. Schoemaker, J. N. H. Reek, P. C. J. Kamer, P. W. N. M. van Leeuwen, M. Lutz and A. L. Spek, *Organometallics*, 2000, **19**, 872.
15 (*a*) P. C. J. Kamer, M. Kranenburg, P. W. N. M. van Leeuwen and J. G. de Vries, *Pat. Appl.* 9400470, 1994, WO95/30680; (*b*) M. Kranenburg, Y. E. M. van der Burgt, P. C. J. Kamer, P. W. N. M. van Leeuwen, K. Goubitz and J. Fraanje, *Organometallics*, 1995, **14**, 3081; (*c*) S. Hillebrand, J. Bruckmann, C. Krüger and M. W. Haenel, *Tetrahedron Lett.*, 1995, **36**, 75; (*d*) P. C. J. Kamer, P. W. N. M. van Leeuwen and J. N. H. Reek, *Acc. Chem. Res.*, 2001, **34**, 895.
16 J. J. Carbó, F. Maseras, C. Bo and P. W. N. M. van Leeuwen, *J. Am. Chem. Soc.*, 2001, **123**, 7630.
17 S. Aguado-Ullate, L. Guasch, M. Urbano-Cuadrado, C. Bo and J. J. Carbó, *Catal. Sci. Technol.*, 2012, **2**, 1694.

18 (*a*) C. Müller and D. Vogt, *Dalton Trans.*, 2007, 5505; (*b*) C. Müller and D. Vogt, in *Phosphorus Chemistry: Catalysis and Material Science Applications*, ed. Peruzzini, M. and Gonsalvi, L., Springer, 2011, ch. 6, vol. 36; (*c*) C. Müller in *Phosphorus Ligand Effects in Homogeneous Catalysis: Design and Synthesis*, ed. P. C. J. Kamer and P. W. N. M. van Leeuwen, Wiley-VCH, 2012; (*d*) L. E. E. Broeckx, M. Lutz, D. Vogt and C. Müller, *Chem. Commun.*, 2011, **47**, 2003; (*e*) C. Müller, L. E. E. Broeckx, I. de Krom and J. J. M. Weemers, *Eur. J. Inorg. Chem.*, 2013, 187.

19 B. Breit, R. Winde and K. Harms, *J. Chem. Soc., Perkin Trans. 1*, 1997, **18**, 2681.

20 (*a*) J. L. Melville and J. D. Hirst, *J. Chem. Inf. Model.*, 2007, **47**, 626; (*b*) B. M. Spowage, C. L. Bruce and J. D. Hirst, *J. Cheminf.*, 2009, **1**, 22.

21 E. Zuidema, E. Daura-Oller, E. J. J. Carbó, C. Bo and P. W. N. M. van Leeuwen, *Organometallics*, 2007, **26**, 2234.

22 (*a*) U. Gellrich, W. Seiche, M. Kelller and B. Breit, *Angew. Chem., Int. Ed.*, 2012, **51**, 11033; (*b*) X. Luo, D. Tang and M. Li, *Int. J. Quantum Chem.*, 2006, **106**, 1844; (*c*) D. Gleich and J. Hutter, *Chem.–Eur. J.*, 2004, **10**, 2435; (*d*) C. R. Landis, *J. Chem. Soc., Dalton Trans.*, 2002, 729; (*e*) S. A. Decker and T. R. Cundari, *Organometallics*, 2001, **20**, 2827; (*f*) S. A. Decker and T. R. Cundari, *J. Organomet. Chem.*, 2001, **635**, 132; (*g*) T. Matsubara, N. Koga, Y. Ding, D. G. Musaev and K. Morokuma, *Organometallics*, 1997, **16**, 1065.

23 A. R. Rossi and R. Hoffmann, *Inorg. Chem.*, 1975, **14**, 365.

24 (*a*) C. Müller, E. A. Pidko, D. Totev, M. Lutz, A. L. Spek, R. A. van Santen and D. Vogt, *Dalton Trans.*, 2007, 5372; (*b*) C. Müller, E. A. Pidko, A. J. P. M. Staring, M. Lutz, A. L. Spek, R. A. van Santen and D. Vogt, *Chem.–Eur. J.*, 2008, **14**, 4899; (*c*) J. J. M. Weemers, W. N. P. van der Graaff, E. A. Pidko, M. Lutz and C. Müller, *Chem.–Eur. J.*, 2013, **19**, 8991.

25 J. J. Carbó, A. Lledós, D. Vogt and C. Bo, *Chem.–Eur. J.*, 2006, **12**, 1457.

26 S. Aguado-Ullate, S. Saureu, L. Guasch and J. J. Carbó, *Chem.–Eur. J.*, 2012, **18**, 995.

27 N. Sakai, S. Mano, K. Nozaki and H. Takaya, *J. Am. Chem. Soc.*, 1993, **115**, 7033.

28 The values of ν_{CO} in the *trans*-$L_2Rh(CO)Cl$ complex are 2016, 2006, 1999, 1979 and 1958 cm^{-1} for L = $P(OPh_3)$, $P(OMe_3)$, 2,4,6-triphenylphosphinine, PPh_3 and PEt_3, respectively. See ref. 18 and the references therein.

29 (*a*) N. S. Antonova, J. J. Carbó and J. M. Poblet, *Organometallics*, 2009, **28**, 4283; (*b*) N. S. Antonova, J. J. Carbó and J. M. Poblet, *Dalton Trans.*, 2011, **40**, 2975.

30 Note that for strongly electron-withdrawing phosphites, the rate-determining step can switch to hydrogenolysis (see ref. 4 and 7). Nevertheless, the comparison could be useful to understand the effects of π-acceptor ligands. Note also that the simplified model ligand has been compared with the real phosphane-phosphite binaphos ligand in the previous work obtaining consistent results for the electronic features (see ref. 26).

31 In the previous contribution we used 3D descriptors, which are more time-consuming and more difficult to interpret, because our main focus was the modeling of enantioselectivity. This property requires 3D descriptors that are able to account for the catalyst shape.

32 (*a*) A. G. Maldonado, J. A. Hageman, S. Mastroianni and G. Rothenberg, *Adv. Synth. Catal.*, 2009, **351**, 387; (*b*) N. Fey and J. N. Harvey, *Coord. Chem. Rev.*, 2009, **253**, 704; (*c*) N. Fey, *Dalton Trans.*, 2010, **39**, 296; (*d*) C. R. Corbeil and N. Moitessier, *J. Mol. Catal. A: Chem.*, 2010, **324**, 146; (*e*) S. Aguado-Ullate, M. Urbano-Cuadrado, I. Villalba, E. Pires, J. I. García, C. Bo and J. J. Carbó, *Chem.–Eur. J.*, 2012, **18**, 14026.

33 (*a*) S. Sciabola, A. Alex, P. D. Higginson, J. C. Mitchell, M. J. Snowden and I. Morao, *J. Org. Chem.*, 2005, **70**, 9025; (*b*) M. Urbano-Cuadrado, J. J. Carbó, A. G. Maldonado and C. Bo, *J. Chem. Inf. Model.*, 2007, **47**, 2228.

34 A. T. Axtell, J. Klosin and A. Abboud, *Organometallics*, 2006, **25**, 5003.

35 J. Gasteiger and M. Marsili, *Tetrahedron*, 1980, **36**, 3219.

36 S. A. Wildman and G. M. Crippen, *J. Chem. Inf. Comput. Sci.*, 1999, **39**, 868.

37 T. J. Hou, K. Xia, W. Zhang and X. J. Xu, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 266.

38 (*a*) *ADF 2007.01.* Department of Theoretical Chemistry. Vrije Universiteit, Amsterdam; (*b*) E. J. Baerends, D. E. Ellis and P. Ros, *Chem. Phys.*, 1973, **2**, 41; (*c*) L. Versluis and T. Ziegler, *J. Chem. Phys.*, 1988, **88**, 322; (*d*) G. Te Velde and E. J. Baerends, *J. Comput. Phys.*, 1992, **99**, 84; (*e*) C. Fonseca Guerra, J. G. Snijders, G. Te Velde and E. J. Baerends, *Theor. Chem. Acc.*, 1998, **99**, 391; (*f*) G. Te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, C. Fonseca Guerra, E. J. Baerends, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931.

39 (*a*) A. D. Becke, *J. Chem. Phys.*, 1986, **84**, 4524; (*b*) A. D. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098.

40 (*a*) J. P. Perdew, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1986, **33**, 8822; (*b*) J. P. Perdew, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1986, **34**, 7406.

41 See for example: (*a*) Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.*, 2007, **3**, 289; (*b*) S. Tsuzuki, *J. Chem. Phys.*, 2001, **114**, 3949; (*c*) E. J. Meijer and M. Sprik, *J. Chem. Phys.*, 1996, **105**, 8684.

42 Z. Yang and D. G. Truhlar, *Acc. Chem. Res.*, 2008, **41**, 157.

43 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787.

44 (*a*) K. Morokuma, *J. Chem. Phys.*, 1971, **55**, 1236; (*b*) T. Ziegler and A. Rauk, *Theor. Chim. Acta*, 1977, **46**, 1; (*c*) T. Ziegler and A. Rauk, *Inorg. Chem.*, 1979, **18**, 1755; (*d*) T. Ziegler and A. Rauk, *Inorg. Chem.*, 1979, **18**, 1558.

45 S. Wold, M. Sjostrom and L. Eriksson, *Chemom. Intell. Lab. Syst.*, 2001, **58**, 109.

46 D. M. Hawkins, S. C. Basak and D. Mills, *J. Chem. Inf. Comput. Sci.*, 2003, **43**, 579.