Coordination Chemistry Reviews 253 (2009) 704–722

Contents lists available at ScienceDirect

Coordination Chemistry Reviews

journal homepage: www.elsevier.com/locate/ccr

Review

Building ligand knowledge bases for organometallic chemistry: Computational description of phosphorus(III)-donor ligands and the metal–phosphorus bond

Natalie Fey*, A. Guy Orpen, Jeremy N. Harvey**

School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

Contents

ARTICLE INFO

Article history:
Received 14 March 2008
Accepted 29 April 2008
Available online 4 May 2008

Keywords:
Ligand effects
Stereoelectronic maps
Computational chemistry
Structure–activity relationships
Phosphorus(III) donor ligands
Data analysis

ABSTRACT

Changing the coordinated ligands is a powerful and synthetically convenient way of modifying and fine-tuning the properties of transition metal complexes, especially those active in homogeneous catalysis. Parameters capturing such changes in the steric and electronic characteristics of complexes have played a key role in improving our understanding of ligand effects on the kinetic, thermodynamic, spectroscopic and structural behaviour of such species. Such ligand parameters can be useful for interpreting experiments, but they can also guide the discovery of novel ligands from ligand maps and allow the prediction of ligand effects before further experimentation. The latter aims especially are best served if such parameters can be determined before ligands and complexes have been synthesised, and here we review calculated descriptors for phosphorus(III) ligands as widely used in organometallic and coordination chemistry. We also discuss the application of such ligand descriptors in models, maps and predictions of ligand effects, describe related computational studies of the metal–phosphorus bond, and provide an overview of the statistical methods used.

1. Introduction

The development, discovery and optimisation of catalytically active transition metal complexes continue to attract considerable research interest; amongst other things, this is often driven by the benefits of achieving efficient and highly selective transformations in synthetic organic chemistry. Finding the best homogeneous

* Corresponding author. Tel.: +44 117 3318260; fax: +44 117 9251295.
** Corresponding author. Tel.: +44 117 9546991; fax: +44 117 9251295.
E-mail addresses: Natalie.Fey@Bristol.ac.uk (N. Fey), Jeremy.Harvey@Bristol.ac.uk (J.N. Harvey).

0010-8545/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.ccr.2008.04.017

organometallic catalyst for a given transformation is a difficult task, as a range of input variables (metal, ligands, co-catalysts, substrates, solvents and reaction conditions including temperature and pressure) need to be optimised. In addition, many reactions in homogeneous catalysis involve a sequence of steps, each of which may be affected in different ways by changing these variables. Perhaps not surprisingly, the "rational design" of new catalysts has rarely been successful and this concept has indeed been criticised recently [1,2]. Instead, the development of new catalysts continues to rely on a combination of chemical intuition and the testing and optimisation of a range of likely candidates. While we thus rarely achieve a truly quantitative appreciation of the effect of changing input variables on the outcome of a reaction (such as activity, chemo-, regio- and enantioselectivity, turnover number and frequency), detailed analysis of experimental results in terms of familiar concepts such as steric and electronic effects can often provide useful insights, in turn guiding further experiments and indeed focussing experimental efforts on the most likely targets.

Changing the ligands coordinated to the metal centre can be a useful way of modifying the properties of transition metal complexes, and such modifications have been studied extensively (see for example Refs. [2–10]). In addition, a range of experimental and calculated parameters have been proposed which seek to capture the steric and electronic properties of common ligand types and hence determine their impact on experimental observations and, more specifically, catalyst properties. Ligands are indeed often the main target in catalyst optimisation, because such property descriptors can be used to guide modifications more efficiently. For these ligands, previous analyses of experimental results have established a *context of knowledge* in which predictions based on these descriptors can be attempted and tested. The interpretation of ligand effects on experimental observables can also help us to identify key properties contributing to a desired result and hence lead to better design criteria. While such models and predictions are currently unlikely to identify a single optimal ligand design for each and every problem, fast calculations allow their application to guide experimental screening towards achieving the most efficient exploration of active ligand space.

Perhaps the most well-established and varied class of ligands are those with phosphorus(III) donor atoms of general structure PA_3, where different substituents can be introduced to modify their steric and electronic properties (A = R, Ar, OR, NR_2, Hal), with further fine-tuning achieved in asymmetric substitution patterns, PA_2A', $PAA'A''$. For this class of ligands the analysis of ligand effects expressed as linear free energy relationships has a long tradition, building on Tolman's seminal work [3] and extended by the QALE approach [11–30] and other work [4,31–35]. However, these studies rely to some extent on experimentally measured descriptors, which cannot necessarily be determined for all known ligands, as there may be experimental limitations, such as solubility, binding and toxicity [36]. In addition, such ligand parameters prevent the consideration of novel ligands, which have not yet been synthesised [36]. More recent work, which will be discussed in this review, has capitalised on the rapid advance of both computing power and computational chemistry techniques to derive suitable descriptors, allowing the theoretical evaluation of both established and novel ligands from parameters which can be calculated quickly and efficiently.

Steric and electronic parameters determined experimentally have been reviewed in the past, for mono- [3,4,25,26,34,37,38] and bidentate [5,39] phosphorus donor ligands and other classes of ligands, *e.g.* *N*-heterocyclic carbenes (NHCs) [9,10,40,41]. In contrast, this review will focus on descriptors for monodentate, phosphorus(III) donor ligands determined by computational means. This will be combined with a discussion of the statistical analysis techniques commonly used to relate both experimental and calculated descriptors with available data (see Refs. [42,43] for more detailed discussions of statistical methods). Our intention is to describe the current state of computational ligand descriptors for phosphorus donor ligands, to summarise computational studies of metal–phosphorus bonding, and to highlight the close link between chemical data and statistical analysis in achieving a more quantitative understanding of ligand effects, leading eventually to successful predictions and hence the true design of new catalysts.

2. Tolman's legacy

While the analysis of substituent effects on experimental data in terms of linear free energy relationships has a long tradition in physical organic chemistry [44], such analyses have also been applied in inorganic and organometallic chemistry; work relevant to P-donor ligands will be discussed throughout this review. Early examples include analyses of P-donor ligand effects on the rates of ligand exchange in terms of relative basicities expressed as half-neutralisation potentials [45] or the related, unit-free pK_a values [46,47]. Such applications of pK_a data have been reviewed extensively (see for example Refs. [4,25,28]), but the analysis of ligand effects in organometallic complexes was influenced more profoundly by Tolman's work, and despite the theoretical focus of this review, a brief summary of the contribution made by Tolman's 1977 review [3] to the analysis of ligand effects will be useful to establish the foundation of many experimental and calculated ligand descriptors relevant to transition metal complexes. As of February 2008 this review has been cited almost 3000 times according to the ISI Web of Knowledge, accumulating around 100 citations every year at a remarkably steady rate. As indicated by the title "Steric Effects of Phosphorus Ligands in Organometallic Chemistry and Homogeneous Catalysis", the main aim of this paper was to draw attention to the relationship between experimental data and ligand steric effects, countering a tradition of analysing experimental data mainly in terms of electronic effects [3,38]. To this end, Tolman defined two parameters, the steric parameter θ and the electronic parameter ν, and published values for a range of ligands in appendices to the review [3].

2.1. Steric parameter, θ

Tolman used a mechanical device and molecular models to measure the ligand cone angle θ for various symmetric ligands. The cone angle is defined as the apex angle of a cylindrical cone centred at a distance of 2.28 Å from the donor atom, corresponding to an average of a range of M–P bond lengths, and extended to touch the van der Waals radii of the outermost atoms. In the molecular models, any substituents were folded back to give a minimum cone angle, which may not be a realistic assumption in sterically hindered systems. However, Tolman discussed potential limitations of this approach, and revised some cone angles to improve agreement with experimental data. Cone angles of unsymmetrically substituted ligands were also defined; these used half angles $\theta_i/2$ determined for each substituent by measuring the angle between the P–M vector (M positioned at 2.28 Å from the P donor atom) and the vector from M to the outermost atom of the substituent. Tolman then assumed that contributions were additive to give:

$$\theta = \frac{2}{3}\sum_{i=1}^{3}\frac{\theta_i}{2} \quad (1)$$

Similarly, for chelating diphosphines, such a half angle $\theta_i/2$ could be determined as the angle between one M–P bond and the vector

bisecting the P–M–P angle. These cone angles have been revised based on crystallographic [34,37,48] and calculated [49] structures as well as detailed analysis of experimental data [26]. In addition, related steric parameters such as the solid angle, which allows for departure from approximate threefold symmetry by calculating the "shadow" of the ligand on a hypothetical sphere [37,50,51], and ligand profiles [52] have been proposed, but Tolman's original values continue to be widely used in essentially their original form.

2.2. Electronic parameter, ν

This parameter, often also abbreviated to TEP (for Tolman electronic parameter), was measured by infrared spectroscopy and corresponds to the vibrational frequency of the A_1 carbonyl stretching mode of $[Ni(CO)_3L]$ complexes in dichloromethane solution, which, according to Tolman [3], can be measured with an accuracy of $\pm 0.3\,cm^{-1}$ for a wide range of ligands. It captures the net electronic effects of the ligands investigated. Using measurements on mixed complexes with PAB_2 substituted ligands to determine individual ligand contributions χ_i, and considering substituent contributions as additive, the TEP could be estimated as

$$\nu = 2056.1 + \sum_{i=1}^{3} \chi_i \quad (2)$$

for additional ligands not probed experimentally. Although this assumption of additivity has been criticised [4,53], net ligand electronic effects continue to be evaluated based on vibrational frequencies of transition metal carbonyl complexes such as ν (see Ref. [38] for a recent review and Ref. [53] for Bartik's more extensive experimental measurements on 157 ligands) or the related ^{13}C NMR chemical shifts of the carbonyl groups, δ [34,54,55].

2.3. Applications of θ and ν

Tolman then explored the use of these parameters in the analysis of multiple experimental datasets, including spectroscopic, kinetic and thermodynamic data [3]. With a view to promoting the cone angle descriptor, he concentrated mainly on identifying steric contributions and two examples of his graphical analyses, illustrating that relationships exist between NMR [56] and kinetic data [57] and the cone angle, have been reproduced in Fig. 1.

In addition, Tolman presented a steric and electronic map of phosphorus donor ligands by plotting ν versus θ (reproduced in adapted form in Fig. 2). He also projected experimental data as a third dimension onto this map to illustrate the contributions of steric and electronic effects in different experimental datasets. Such three-dimensional plots are shown in Fig. 3 for ^{13}C chemical shifts in $[Ni(CO)_3L]$ complexes [58] and in Fig. 4 for the enthalpies of ligand exchange in a platinum complex (*trans*-$[MePt(PMe_2Ph)_2(thf)]^+ + L \rightarrow$ *trans*-$[MePt(PMe_2Ph)_2L]^+ + thf$) [59]. The related two-dimensional scatter plots highlight perhaps more clearly that the NMR data correlate quite highly with the electronic parameter derived from the same complex (Fig. 3), while the enthalpies are related more closely to the steric parameter (Fig. 4).

We have discussed Tolman's results at length here, because his work has had such a lasting impact on the area of quantitative studies of ligand effects. The cone angle and electronic parameter have become very common measures of ligand properties reflected in the ever-increasing number of citations of Tolman's review papers. Also, the notion of a 'ligand map' (Fig. 2) has become of widespread use and is certainly a guiding principle for many people when considering ligand effects. Indeed, many of the newer computational methods described below are used (among other applications) to generate ligand maps very similar in concept to that of Tolman (Section 5.2).

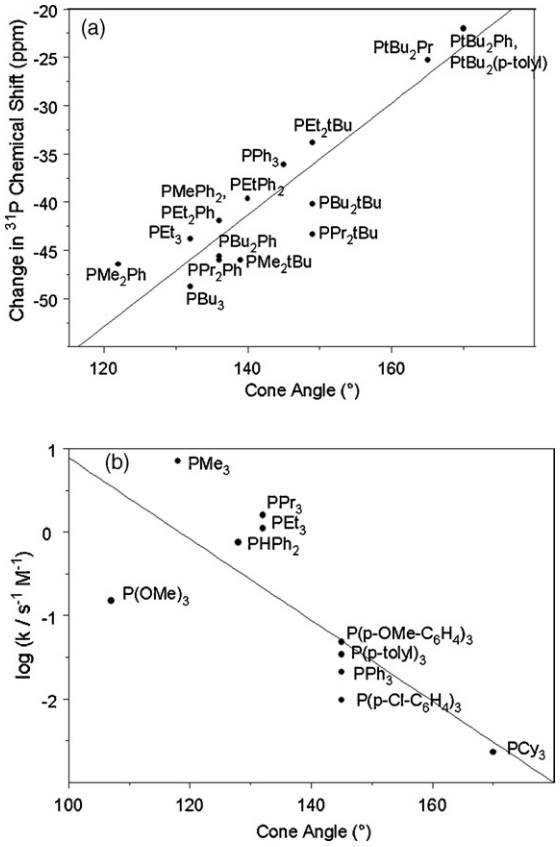


Fig. 1. Correlations with cone angles θ (adapted from Ref. [3]): (a) change in phosphorus chemical shifts on coordination ($\delta_{complex} - \delta_{free}$) from ^{31}P NMR spectra of *trans*-$[RhCl(CO)L_2]$ complexes, L as shown, Ref. [56], and (b) rate constants for the reaction $2Co(DH)_2L + PhCH_2Br \rightarrow PhCH_2Co(DH)_2L + BrCo(DH)_2L$ in benzene, where DH = dimethylglyoxime or 1,2-cyclohexanedione dioxime, L as shown, Ref. [57].

3. Calculated descriptors

This section will focus on individual descriptors of ligand steric and electronic properties which have been derived from calculations and which seek to separate steric and electronic effects.

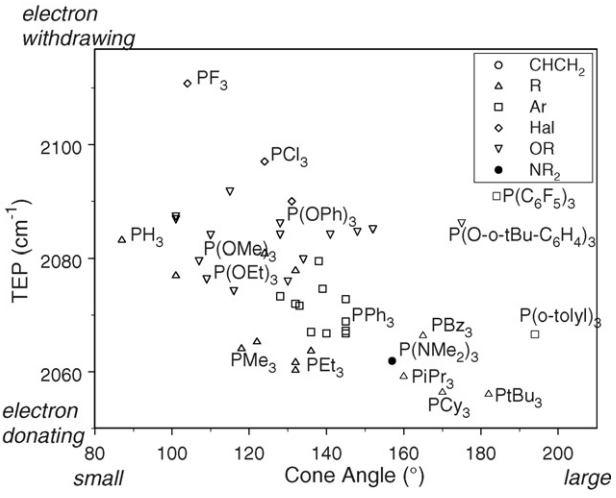


Fig. 2. Tolman map (adapted from Ref. [3]): scatter plot of TEP ν vs. cone angle θ.

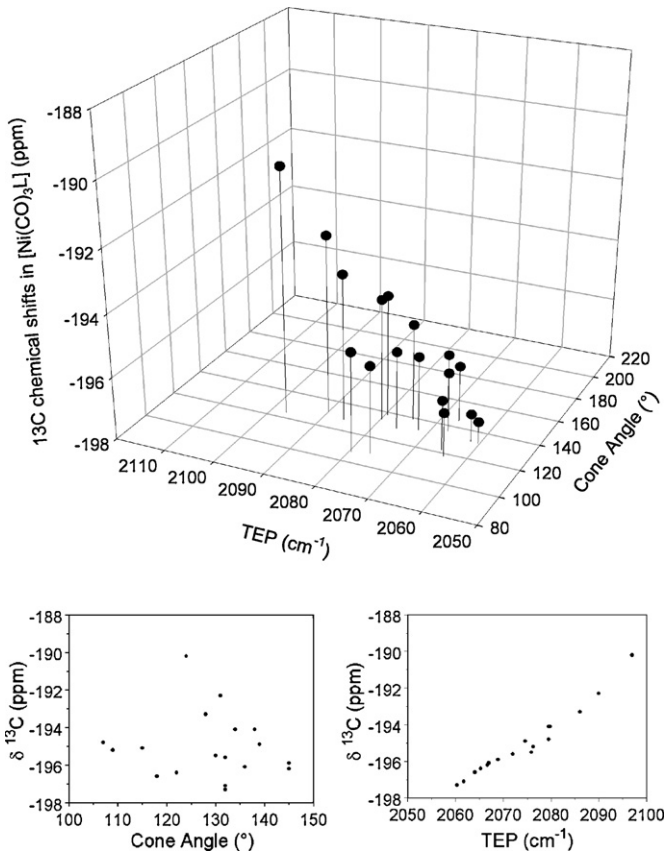


Fig. 3. Stereoelectronic map and scatter plots relating cone angle θ and TEP ν to ^{13}C chemical shifts in $[Ni(CO)_3L]$ complexes [58] (plots adapted from Ref. [3]).

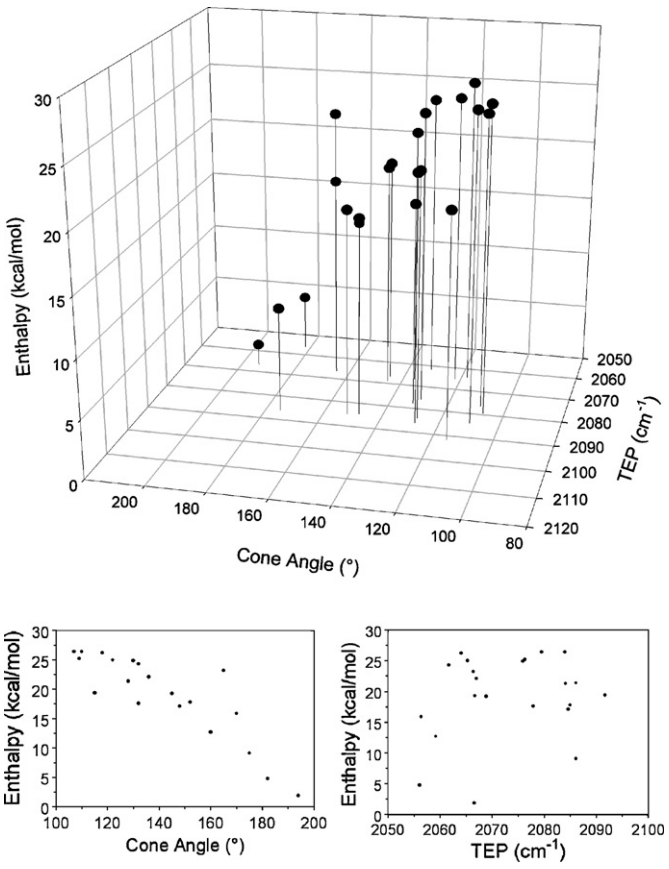


Fig. 4. Stereoelectronic map and scatter plots relating cone angle θ and TEP ν to $-\Delta H$ of the reaction *trans*-$[MePt(PMe_2Ph)_2(thf)]^+ + L \rightarrow$ *trans*-$[MePt(PMe_2Ph)_2L]^+ + thf$ [59] (plots adapted from Ref. [3]).

Approaches aimed at providing a more detailed decomposition and analysis of electronic effects will be discussed in Section 4 and databases of calculated ligand descriptors and their application in the interpretation and prediction of ligand effects will be considered in Section 5 below.

Work reporting individual calculated descriptors often seeks to establish a relationship between this new descriptor and more familiar parameters, such as the cone angle θ, by reporting correlation and regression coefficients. These measure how well such a relationship can be described by a simple linear equation ($y = ax + b$), or indeed a more complicated, non-linear equation [42,43]. The quality of fit to the equation is expressed by linear bivariate correlation (R) and linear regression (R^2) coefficients, with values close to 0 indicating a poor fit while values close to 1 show a good fit. Note that parameters can be inversely correlated, where a good fit would show values of R close to -1. For example, the relationship between ^{31}P NMR data [58] and Tolman's cone angle [3] illustrated in Fig. 1a is described well by a linear equation, giving a high bivariate linear correlation coefficient $R = 0.926$, while cone angle and the enthalpies of ligand exchange for a platinum complex [59], shown in Fig. 4, are inversely correlated with $R = -0.868$. Correlation coefficients only explore whether such a relationship exists, while regression analysis also allows to derive the coefficients of variables in the underlying equation, and hence to make predictions. In addition, regression analysis can be performed for several variables (multivariate linear regression, see Section 5), whereas simple correlation only considers the relationship between two variables. As illustrated by Tolman and in Figs. 1, 3 and 4 above, such descriptor relationships can also be explored and illustrated by appropriate scatter plots, which often support discussions of correlation and regression coefficients in published work.

3.1. Ligand size

Computational steric parameters reported before 1993/1994 have been covered in reviews by Brown and Lee [34] and White and Coville [37]. Much of this early work was directed at comparing cone and/or solid angles for different conformers of individual ligands, optimised with molecular mechanics or semi-empirical approaches. Such work recognised that Tolman's approach [3] of minimising cone angles by folding all substituents inwards (discussed above) is often inappropriate when compared to calculated and crystallographically observed ligand geometries, especially for sterically demanding and/or conformationally flexible ligands. However, reliably identifying the lowest energy conformer for *each* reaction environment remains difficult to this date and implies extensive database mining and conformational searches for every ligand. In addition, even a cone angle determined computationally or crystallographically from a low energy conformer may not reflect the conformer of a metal–ligand complex present during its reactions (see for example Refs. [60,61] for a recent exploration of conformational changes in palladium catalysed cross-coupling). In early studies of ligand size the role of computation was generally limited to providing ligand and transition metal complex geometries for cone/solid angle measurements, along with approximate energetic weightings [51,62]. They focussed on ligand conformational preferences, but are not conceptually different from cone and solid angle measurements on models or crystal structures and will not be discussed further. More recently, Senn and co-workers have also calculated cone angles from optimised geometries and compared them with Tolman's results [49]. This work formed part of an

exploration of the properties of 8 P-donor ligands, mainly aimed at determining differences in the electron donation (see below).

In a series of papers Brown and co-workers developed a molecular mechanics measure of ligand steric effects, the ligand repulsive energy, E_R [34,54,63–65]. This energy can be calculated in three steps: (1) the global minimum energy conformer of the complex is located in a conformational search, using a suitable force field [63], (2) the van der Waals potential function of the force field is then replaced by a purely repulsive form, and (3) a fixed geometry scan of Cr–L distances is performed [54]. The repulsive energy E_R is calculated from the gradient of the repulsive potential function with respect to the Cr–L distance, at the bond length corresponding to the optimised, equilibrium Cr–L distance obtained using the normal force field. These energies correlate reasonably well with cone angles ($R = 0.877$). They were determined for 60 P-donor ligands in Brown's original work [34,54] and have been used in the derivation of linear free energy relationships with experimental data (discussed below, Section 5). Bubel and co-workers later reported an automated approach to the calculation of the ligand repulsive energy, ERCODE, along with values for 112 P-donor and several other classes of ligands [66].

A structural measure of steric bulk for P-donor ligands, the S4′ parameter, which was originally introduced by one of us for the analysis of crystallographic data [67], has been adapted as a computational ligand descriptor by the group of Cundari [68]. This descriptor is calculated as the difference between ∠M–P–A and ∠A–P–A angles in metal complexes of substituted P-donor ligands

$$S4' = \sum(\angle M\text{–}P\text{–}A) - \sum(\angle A\text{–}P\text{–}A) \quad (3)$$

and is inversely correlated with ligand steric bulk, *i.e.* large values of S4′ correspond to small ligands. It captures structural changes in response to increased steric bulk, as the steric interactions between substituents force a larger ∠A–P–A angle and hence a smaller ∠M–P–A angle. In Cundari's work [68], these angles were taken from optimised geometries of a series of rhodium complexes [*trans*-$RhL_2(CO)Cl$] obtained with the semi-empirical PM3(tm) approach after extensive molecular mechanics conformational searches aimed at locating the global minimum. Although the crystallographic version of S4′ [67] (S4′ crystal) considers a range of metal and non-metal phosphine complexes, agreement with the calculated data (S4′ Rh calc.) is generally good for a set of 16 ligands common to both studies (regression coefficient $R^2 = 0.957$). This parameter has been calculated for 95 phosphines and phosphites, including a range of exotic ligands where the donor atom is part of a ring system. These data were combined with a calculated electronic descriptor (SEP, discussed below) to develop a map of ligand space (see Section 5.2, Fig. 9).

Suresh has reported a steric descriptor derived from quantum mechanical calculations, which essentially captures the same structural changes as S4′, but only considers the free ligand [69]. This parameter is an extension of the molecular electrostatic potential (MESP) approach developed as a P-donor ligand electronic parameter [70] (discussed in Section 3.2 below). It is calculated from a hybrid quantum mechanics/molecular mechanics QM/MM calculation using the ONIOM method [71], where a PH_3 group is used to represent the ligands in the QM section and any substituents are treated with MM; the layers are connected by hydrogen link atoms. The shape of the electrostatic potential created by each ligand is affected by steric interactions between the substituents, which force larger ∠H–P–H angles and increase the p-orbital contribution to the lone pair on the sp^3-hybridised phosphorus atom. As only PH_3 is considered quantum mechanically, substituent electronic effects are not captured in the layered model and any geometry changes are due to steric interactions between substituents in the MM region. A local minimum of the electrostatic potential created by the ligand, corresponding roughly to the local maximum in the lone pair electron density, can be located from visual inspection of the MESP function to give V_{min}. The steric descriptor can then be calculated as the change in V_{min} for the different ligands PR_3 compared to free PH_3

$$MESP_{steric} = V_{min}(PH_3) - V_{min}(ONIOM, PR_3) \quad (5)$$

This parameter has been calculated for 21 alkyl- and arylphosphine ligands and Suresh showed that it is highly correlated with Tolman's cone angles ($R = 0.976$).

In later work [72], Suresh and co-workers extended this approach to explore the steric properties of 46 ligands, including exotic ligand structures identified from the Cambridge Structural Database (CSD) [73]. They also used DFT-optimised geometries to calculate S4′-type data for the free ligands (DFT-S4) and compared these values with the corresponding data determined from crystal structure geometries ($R = 0.889$). Their discussion explored the impact of electronic contributions from the coordination environment and the substituents on the ligand geometry, which can be isolated in appropriate scatter plots. ONIOM-S4 data (Table 1), which they described as free of electronic effects from substituents, correlate highly with S_{eff} data ($R = 0.921$). (While the definition for S_{eff} is the same as for $MESP_{steric}$, Eq. (5), the sign of the data has changed for alkylphosphine ligands present in both sets and S_{eff} is thus positive.) By combining the parameter with a measure of electronic effects, E_{eff}, derived from the same calculations (see Section 3.2), they also generated a map of phosphine ligands (discussed in Section 5.2).

We have recently described the development of a ligand knowledge base (LKB) of calculated descriptors for phosphorus(III) donor ligands, LKB-P [74,75]. This database collects structural, energetic and charge descriptors (see Table 3 below) calculated with density functional theory (DFT) for a range of ligands, both free and in representative complexes, and we have illustrated the design, potential applications and suitable analysis techniques of such a LKB [74,75]. The steric properties of ligands likely contribute to a number of LKB descriptors, but we have also included two descriptors specifically to capture ligand size: the S4′ parameter described above and in this case calculated from the relevant structural parameters in the DFT-optimised complexes BH_3.L, $[PdCl_3L]^-$ and $[Pt(PH_3)_3L]$ (LKB S4′), and a new steric parameter, the He_8_steric interaction energy. The He_8_steric descriptor is designed to mimic the non-bonded interactions between the P-donor ligand and other *cis* ligands in an octahedral complex and it is calculated as the energy of interaction between the ligand and a fixed ring of eight helium atoms

Table 1
Linear bivariate correlation coefficients *R* for steric descriptors (excl. P(o-tolyl)$_3$ and P(C$_6$F$_5$)$_3$, see text)

	Cone angle [3] (°)	E_R [34,54] (kcal mol^{-1})	ERCODE [66] (kcal mol^{-1})	S4′ crystal [67] (°)	S4′ Rh calc. [68] (°)	$MESP_{steric}$ [69] (kcal mol^{-1})	S_{eff} [72] (kcal mol^{-1})	DFT-S4, ligand [72] (°)	ONIOM-S4 [72] (°)
Number of ligands	23	25	26	14	18	14	13	13	13
R for linear fit vs. He_8_steric	0.877	0.956	0.844	−0.758	−0.740	−0.834	0.891	−0.745	−0.867
R for linear fit vs. LKB S4′	−0.773	−0.798	−0.796	0.881	0.778	0.718	−0.797	0.972	0.783

with a radius of 2.5 Å. The ligand phosphorus atom is constrained to lie exactly 2.28 Å above the centroid of this ring, but the ligand geometry itself is re-optimised, starting from an optimised conformation of the free ligand. Data for 61 P-donor ligands have been published [74].

In our original work [74] we established the relationship between our He_8_steric descriptor and cone angles. These descriptors are correlated, but deviations from a linear relationship occur for bulky ligands where different conformers may need to be considered. Fig. 5 explores the relationship between the He8_steric descriptors and some of the steric descriptors discussed here.

As observed in our original analysis [74], our He_8_steric descriptor and the cone angles reported by Tolman [3] have a reasonably high linear correlation coefficient; here we have removed the P(o-tolyl)$_3$ and P(C_6F_5)$_3$ ligands from consideration (see [74] for a more detailed discussion of conformational issues) to give an improved fit to a linear relationship (R = 0.877, see Table 1). However, the relationship would be slightly better described by a cubic function as indicated (regression coefficient, R^2 = 0.816 versus 0.770 for linear fit), which avoids unphysical predictions (<0) for the He_8_steric parameter where ligands are small and takes better account of the observed deviations for large ligands. These deviations likely arise because the optimised ligand geometries do not necessarily minimise ligand size, whereas Tolman folded all substituents away from the metal in order to achieve the smallest possible cone angle.

Optimised low energy ligand conformers were used for all of the calculated descriptors discussed here and these showed quite high linear correlations (Fig. 5b–d, Table 1) with He_8_steric. Differences between E_R and ERCODE presumably arise from different conformational search results, giving, *e.g.* remarkably different values for PCy_3 ligands (E_R = 116, ERCODE = 66) which may indeed adopt a range of different conformations, but should be quite similar to P^iPr_3 (E_R = 109 and ERCODE = 107) [34,54,66]. Correlations between He_8_steric and S4′ parameters determined by different approaches vary, but are reasonably high, and the variations in data perhaps reflect the different ligand environments, with the crystallographic [67] and LKB [74] versions of the descriptor collated from a range of coordination complexes while Cundari's parameter [68] considered a single rhodium complex and Suresh's data [72] were derived from free ligands. The differences between S4′ parameters of Suresh et al. can be related to the electronic effects of substituents, as discussed in their work [72], which affect DFT-S4 but are absent from ONIOM-S4. The correlation between the He_8_steric and LKB S4′ descriptors is reasonably high (R = −0.701) and significant linear correlations can also be observed between LKB S4′ and the other steric descriptors (Table 1).

Deviations from a linear relationship between these steric parameters and our He_8_steric descriptor are often more pronounced for small ligands, where the He_8_steric data converge to a minimum corresponding to the repulsion between the phosphorus lone pair and the He_8 ring. Several small ligands are near this value, but most other steric descriptors continue to measure the size of ligands with no limiting value, giving a more pronounced differentiation of ligands when these are small. However, in some case the optimised metal–ligand distance is implicit in the descriptor (E_R and ERCODE) or affects the coordination geometry (S4′), and this may respond to changes in the electronic structure of ligands, hence altering the *steric* measure. By including both the ligand-derived He_8_steric and the S4′ descriptor calculated from a range of coordination environments in our ligand knowledge base, we have

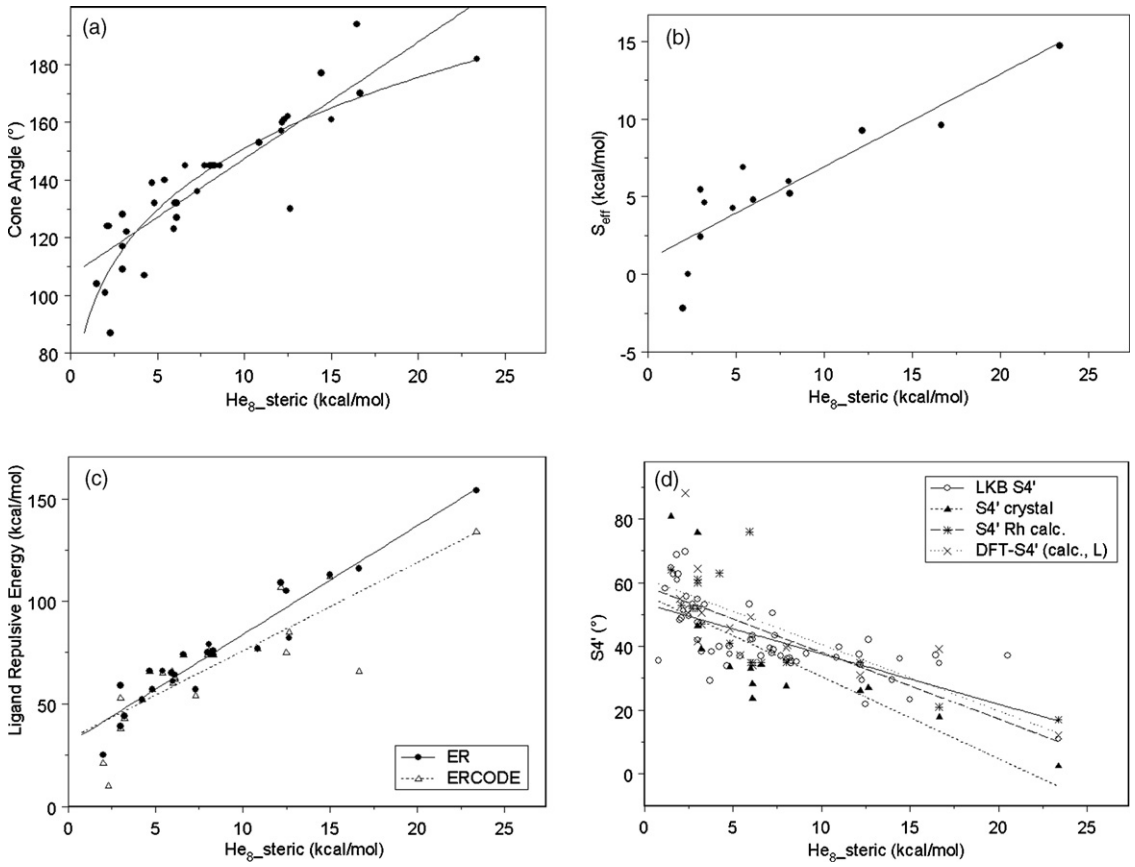


Fig. 5. Relationship between LKB He_8_steric [74] and steric descriptors (excl. P(o-tolyl)$_3$ and P(C_6F_5)$_3$, see text): (a) cone angle θ [3]; (b) S_{eff} [72]; (c) E_R [34,54] and ERCODE [66]; (d) S4′ derived from various sources (see text and Table 1 for details) [67,68,72,74].

sought to take account of both effects, thereby optimising transferability of the database to a variety of chemical environments (see below).

While Tolman's concept of a cone angle [3] remains perhaps the most widely used measure of ligand steric bulk, descriptors proposed more recently allow one to take better account of a ligand's conformational responsiveness to changes in the coordination environment. Rather than relying on a single "compressed" ligand geometry, such descriptors can potentially be calculated/determined from a range of different conformations to capture the more dynamic aspects of transition metal complexes. It is perhaps worth noting, though, that most descriptors are derived from a single ligand conformer, albeit often the relevant global minimum, and one might speculate on the impact of conformational noise on such descriptor values [74,76]. For monodentate P-donor ligands, such effects are likely to be small for most ligands, but a few pathological cases (see discussion above and Ref. [74]) should be expected.

Combining calculated structural data with an exploration of conformational preferences, whether for cone angles or other descriptors, allows the consideration of novel ligand architectures before synthesis. However, the main appeal of cone angles over other steric descriptors remains their reassuring familiarity (and their linkage with experimental observables), and we expect that some of the descriptors discussed here, which can be calculated in a straightforward and more reliable fashion, will become established as useful alternatives.

3.2. Ligand electronic properties

Computational studies of the electronic properties of P-donor ligands crucially depend on the availability of suitable theoretical approaches and resources, and few studies reported before 2000 consider large samples of ligands; these will not be discussed further here.

In a study of eight P-donor ligands, Senn and co-workers explored the use of DFT-calculated HOMO energies (E_{HOMO}) for the free ligands as measures of donor strength [49]. They also derived cone angles from their optimised geometries (see comments above). Observing a reasonably linear relationship between the E_{HOMO} descriptor and both experimental and calculated proton affinity data on a scatter plot, they further investigated the correlation of E_{HOMO} with a calculated activation energy $\Delta E^{\ddagger}$ for the protolytic cleavage of the nickel-carbon bond in $[NiCl(CH_2CH_2NH_3)L_2]^+$. This is quite high and a regression coefficient $R^2 = 0.976$ could be achieved for a simple linear model, which allows the quantitative prediction of activation barriers for other ligands where steric properties have not changed (*para*-substituted aryl phosphines).

Despite the widespread use of Tolman's electronic parameter [3] discussed above, in some cases the TEP and related carbonyl stretching frequencies are of limited use as electronic descriptors, due to the fact that the metal carbonyl complexes of some interesting ligands are too toxic or unstable to be readily studied, or because such ligands cannot be (or have not yet been) prepared [36]. Such issues were addressed by Clot and co-workers when they presented a computationally derived ligand electronic parameter (CEP) [36], based on DFT-calculated vibrational frequencies in the familiar $[Ni(CO)_3L]$ complex proposed by Tolman [3]. They reported values for 68 ligands, of which 12 were P-donor ligands, and established correlations between their CEP and various experimental ligand electronic descriptors, such as the TEP [3] ($R = 0.982$), the Hammett substituent constant σ_m [44] ($R = 0.967$) and the electrochemical parameter E_0 developed by Lever (LEP) [77,78] ($R = 0.960$). Using linear regression equations derived in their work, it was shown that the CEP could also be used to estimate values for these other electronic descriptors.

As frequency calculations with DFT approaches can be computationally expensive, Cundari, White and co-workers explored the use of a semi-empirical (PM3(tm)) approach for the calculation of vibrational frequencies in suitable carbonyl complexes [79]. They initially compared the A_1 carbonyl vibrational frequencies calculated for two different complexes ($[Mo(CO)_5L]$, $[W(CO)_5L]$) and the carbonyl stretching frequency of $[Rh(Cp)(CO)L]$ with established electronic parameters: TEP [3] ($R = 0.952$ (Mo), 0.951 (W), 0.913 (Rh)), LEP [77,78] ($R = 0.960$ (Mo)) and CEP [36] ($R = 0.954$ (Mo)), and concluded that their semi-empirical electronic parameters (SEP) are robust, fragment-independent measures of electronic effects. They report the SEP_Mo for 24 ligands, including 15 P-donor ligands and both SEP_W and SEP_RhCp for 46 ligands, of which 37 have phosphorus donor atoms. In later work, they applied the same protocol to [*trans*-$RhL_2(CO)Cl$] after MM conformational searching and semi-empirical geometry optimisations as described above for the semi-empirical version of S4′ Rh [68]. In this case, the descriptor SEP_Rh has been reported for 95 phosphines and phosphites, including a range of exotic ligands where the donor atom is part of a ring system.

Suresh's molecular electrostatic potential (MESP) approach described above was originally reported as a measure of the electronic effects in substituted P-donor ligands [70]. For the electronic parameter, the full ligand is optimised at the DFT level and the local minimum of the electrostatic potential created by the ligand in the lone pair region, V_{min}, is derived from the computed electron density. This descriptor has been calculated for 33 phosphine and phosphite ligands and related to the TEP [3] ($R = 0.973$), as well as the pK_a values of protonated ligands ($[PR_3H]^+$ [46], $R = 0.960$) and several calculated and experimental reaction energies. These included for example the calculated energy released in the formation of $[Ni(CO)_3L]$ from $[Ni(CO)_3]$ and free L (ΔE, $R = 0.979$) and the experimental enthalpy (ΔH°, $R = 0.984$) and the reduction potentials (E°, $R = 0.985$) for the electrochemical reaction $[\eta\text{-Cp(CO)L(COMe)Fe}]^+ + e^- \rightarrow [\eta\text{-Cp(CO)L(COMe)Fe}]$ [17]. The authors also explored the relationship between V_{min} and the phosphorus lone pair energy expressed as E_{HOMO}, proposed by Senn et al. [49]. Correlation was high for alkylphosphines ($R = 0.993$), but less good for other ligand types. This was related to poorer correlations between E_{HOMO} and the reaction energies considered (*e.g.* for ΔE, $R = 0.848$, for ΔH°, $R = 0.753$ and for E°, $R = 0.752$) and Suresh et al. thus proposed V_{min} as a superior descriptor for the electron donor properties of P-donor ligands. The validity of their analyses was later questioned by Giering, Prock and Fernandez, who suggested that the observed linear correlations might be due to other factors, and that not enough ligands had been considered in the original study to be able to detect such effects [22]. However, Kühl's recent review is more favourable, citing the high correlation with TEP [38].

In more recent work aimed at exploring both the steric and electronic effects of ligands [72], Suresh and co-workers derived an electronic parameter (E_{eff}) from two-layer QM-MM calculations (described above) of V_{min}

$$E_{eff} = V_{min}(ONIOM\_PR_3) - V_{min}(PR_3) \quad (6)$$

This electronic parameter has been calculated for 46 ligands, but, with the work focussing on the discussion of steric effects (Section 3.1) and the generation of a ligand map (Section 5.2), was not discussed in detail.

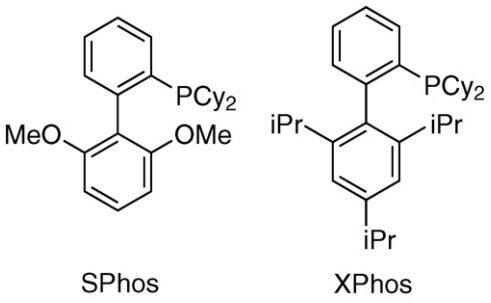
Our ligand knowledge base [74,75] also contains a range of electronic descriptors (see Table 3 below), including a DFT-calculated proton affinity (PA) and the frontier molecular orbital energies (E_{HOMO}, E_{LUMO}), which correlate with analogous descriptors

reported by other groups (Table 2). Correlation with the σ bonding descriptors (PA, E_{HOMO}) becomes poorer for parameters derived from metal complexes (TEP, CEP, SEP), where π back-bonding may occur (these show some correlation with E_{LUMO}) and interactions with other ligands can affect the electron density. Interestingly, correlations between V_{min} and the LKB σ bonding descriptors are quite low, even though V_{min} should also capture σ bonding effects. Suresh and Koga explored the correlation between the energy of the phosphorus lone pair orbital (equivalent to E_{HOMO}) and their V_{min} descriptor [70] and found good agreement for PA_3 ligands (A = H, alkyl, R = 0.993), but a wider scatter of points for other substituents, suggesting perhaps that the two descriptors capture different aspects of σ bonding.

Although the LKB descriptors E_{HOMO} and PA can be considered σ bonding descriptors and E_{LUMO} may be related to π back-bonding, we have not attempted to separate steric and σ/π electronic effects explicitly, choosing instead to extract structural and energetic descriptors from a range of metal and Lewis acid fragments ($[HL]^+$, BH_3.L, $[PdCl_3L]^-$ and $[Pt(PH_3)_3L]$, see Table 3 below) to capture variations in metal–ligand bonding. As well as E_{HOMO}, E_{LUMO} and PA discussed above, these descriptors include adduct binding energies, ligand and metal fragment charges from natural bonding orbital (NBO) analysis, structural parameters capturing geometry changes of both ligand and metal fragments upon complexation and the steric descriptors discussed in Section 3.1. Few of these descriptors are purely electronic, as steric interactions with bulky ligands can affect the complex geometries, but many show high correlations with the parameters discussed here (*e.g.* between TEP and BE(B) $R = 0.947$; TEP and <$(H_3P)Pt(PH_3)$ $R = 0.988$; SEP_Rh and BE(Pt) $R = 0.909$; SEP_Mo and P-Pt $R = -0.992$, see Table 3 for LKB descriptor names), suggesting that they capture similar ligand properties. By considering a range of descriptors derived from different coordination environments, we explore the responsiveness of ligand electronic properties, which may remain obscured if only a single electronic descriptor is considered. In addition, we have demonstrated how these descriptors can be used to derive good linear regression models for several experimental datasets [74,75], including the TEP [3], thus demonstrating the excellent transferability of descriptors to a variety of chemical environments. Section 4 will consider how different contributions from σ and π bonding or other effects could be identified computationally and the application of the LKB in maps and models of ligand space will be discussed in detail in Sections 5.2 and 5.3 below.

4. Computational analysis of metal–phosphine bonding

The methods for computational determination of ligand descriptors discussed in the previous section mostly attempt to calculate known experimental observables, or related properties that correlate strongly with such data. In this sense, they often assume that ligand–metal interactions are well characterized by just two parameters, namely steric effects, measured by cone angles, and electronic effects, measured by the TEP. The usefulness of a computational descriptor tends to be considered as being highest in this framework when it provides a good correlation with the original Tolman parameters or their updated equivalents. As suggested in some of the discussion above, however, some important properties of ligands are not sufficiently well resolved by the Tolman parameters, so that successful computational descriptors should in fact go beyond these two to achieve a fuller description of the metal–ligand interactions. In this section, we review methods that have been used to attempt to provide a more thorough description of metal–ligand binding based on first principles. For other reviews on this topic, see Ref. [80].



Scheme 1. Examples of biaryl ligands used in amination.

Steric bulk is indubitably an important property of ligands, and the cone angle captures it reasonably well. However, the detailed effect of ligand sterics will depend on the anisotropy of the steric bulk, and this needs to be measured using more than one parameter [37,50–52]. Also, conformational changes can tune the steric bulk of a given ligand, so that it appears to change size during a given reaction. In some cases, such changes will not be possible, as binding of the ligand to any of the metal species involved in the mechanism will require that it adopts the least bulky of its energetically viable conformations. However, in other cases, a ligand may change conformation during the reaction, adopting a more bulky geometry and thereby 'protecting' the metal centre or a metal-bound functional group in some intermediates while 'folding' itself into a more compact geometry to allow for closer access of other ligands or substrates to the metal centre in others.

An example of this behaviour apparently occurs for some of the bulky diaryl ligands (Scheme 1) developed by Buchwald and others for use in palladium catalysed amination reactions (see Ref. [7] for some recent reviews of cross-coupling reactions). Computational studies of the catalytic cycle with two ligands used in experimental studies, SPhos and XPhos (Scheme 1) [60,61], have shown that in the catalytic cycle low-coordinate palladium complexes are stabilised by secondary interactions with the ligand [60] while favourable amine binding requires rotation of the ligand to remove steric bulk from the metal centre [61] (Fig. 6). The use of a single descriptor to measure the steric bulk of such a ligand would either predict it to be too small or too large, depending on the conformation chosen. And in each case, this would lead to an incorrect conclusion, suggesting that other, less conformationally flexible ligands, of similar bulk to either of the two conformations, would be equally successful in catalysis. The computational exploration of such conformational changes for large ligands in catalytic cycles remains challenging,

Table 2
Linear bivariate correlation coefficients *R* for electronic descriptors

	TEP [3] (cm^{-1})	E_{HOMO} [49] (eV)	PA [49] ($kJ\,mol^{-1}$)	CEP [36] (cm^{-1})	SEP_Mo [79] (cm^{-1})	SEP_W [79] (cm^{-1})	SEP_RhCp [79] (cm^{-1})	SEP_Rh [68] (cm^{-1})	V_{min} [70] ($kcal\,mol^{-1}$)	E_{eff} [72] ($kcal\,mol^{-1}$)
Number of ligands	32	8	8	10	13	28	28	20	19	13
R for linear fit vs. E_{HOMO}	−0.916	0.948	0.913	−0.936	−0.734	−0.842	−0.680	−0.768	−0.572	0.760
R for linear fit vs. E_{LUMO}	−0.444	0.264	0.087	−0.683	−0.596	−0.534	−0.497	−0.476	−0.543	0.569
R for linear fit vs. PA	−0.932	0.920	0.918	−0.950	−0.730	−0.794	−0.598	−0.741	−0.627	0.661

Table 3
Descriptors in LKB-P [74,75]

Descriptor	Derivation (Unit)
Free phosphorus(III) species (L)	
E_{HOMO}	Energy of highest occupied molecular orbital (Hartree)
E_{LUMO}	Energy of lowest unoccupied molecular orbital (Hartree)
LP s-character	Contribution of P s-orbital to lone pair (LP), from NBO analysis (%)
He_8_steric	Interaction energy between L in ground state conformation and ring of eight helium atoms, $E_{ster} = E_{tot}(system) - [E_{tot}(He_8) + E_{tot}(L)]$ (kcal mol^{-1})
Protonated Ligand ($[HL]^+$)	
PA	Proton affinity, calculated as the difference between the energy of the neutral and protonated L (kcal mol^{-1})
Borane adduct ($H_3B.L$)	
Q(B fragm.)	NBO charge on BH_3 fragment
BE(B)	Bond energy for dissociation of P-ligand from BH_3 fragment, BE = $[E_{tot}(fragment) + E_{tot}(L)] - E_{tot}$ (complex) (kcal mol^{-1})
ΔP–A(B)	Change in average P–A bond length compared to free ligand (Å)
ΔA–P–A(B)	Change in average A–P–A angle compared to free ligand (°)
P–B	P–B distance (Å)
Palladium complexes ($[PdCl_3L]^-$	
Q(Pd fragm.)	NBO charge on $[PdCl_3]^-$ fragment
BE(Pd)	Bond energy for dissociation of L from $[PdCl_3]^-$ fragment, BE = $[E_{tot}(fragment) + E_{tot}(L)] - E_{tot}(complex)$ (kcal mol^{-1})
ΔP–A(Pd)	Change in average P–A bond length compared to free ligand (Å)
ΔA–P–A(Pd)	Change in average A–P–A angle compared to free ligand (°)
P–Pd	P–Pd distance (Å)
Pd–Cl trans	Pd–Cl distance, *trans* to L (Å)
Platinum complexes ($[Pt(PH_3)_3L]$)	
Q(Pt fragm.)	NBO charge on $[(PH_3)_3Pt]$ fragment
BE(Pt)	Bond energy for dissociation of P-ligand from $[Pt(PH_3)_3]$ fragment, BE = $[E_{tot}(fragment) + E_{tot}(L)] - E_{tot}(complex)$ (kcal mol^{-1})
ΔP–A(Pt)	Change in average P–A bond length compared to free ligand (Å)
ΔA–P–A(Pt)	Change in average A–P–A angle compared to free ligand (°)
P–Pt	P–Pt distance (Å)
∠$(H_3P)Pt(PH_3)$	Average $(H_3P)Pt(PH_3)$ angle (°)
Cumulative	
S4′ calc	$(\sum <ZPA - \sum <APA)$, where Z = BH_3, $[PdCl_3]^-$, $[Pt(PH_3)_3]$ (°)

as the quantum mechanical and DFT approaches required to yield reliable reaction energies are too expensive for detailed conformational searches, while agreement with relative conformer energies derived from other approaches (molecular mechanics and semi-empirical) can be poor. However, steric descriptors derived from a range of different coordination environments may provide an indication of where such explorations would be useful, focussing our efforts on ligands able to show such a conformational response.

For electronic effects, it is equally clear that the TEP, or any other single descriptor, cannot adequately capture all the electronic properties of a given ligand. For example, two ligands may be roughly equally electron donating in a $Ni(CO)_3(L)$ complex, but one of them may be much more electron donating in an Fe(II) or Rh(III) complex; ligand electronic properties should thus be considered as responsive to changes in their coordination environment. The most obvious reason why this behaviour may arise is that bonding is usually accepted to involve both σ donation of the phosporus lone pair into an empty orbital on the metal, and π back-donation from filled d orbitals on the metal to empty orbitals of appropriate π symmetry

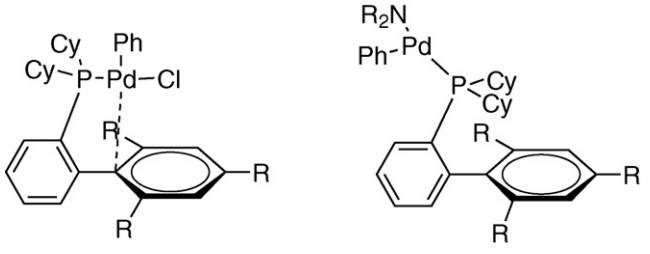


Fig. 6. Different ligand conformations observed in palladium-catalysed amination [60,61].

with respect to the metal–P bond on the ligand [81,82]. The latter orbitals are now generally accepted to be combinations of P–R σ* orbitals [83], although earlier work speculated on the role of vacant phosphorus d orbitals for π acceptance (see for example Ref. [84] for a review). While bonding in some complexes may involve relatively little back-bonding, in others it can play a much larger role, and this can lead to a breakdown of models relying on a single electronic parameter.

Computational results can be used to derive more sophisticated models of metal–ligand bonding, in which the individual contributions of effects such as σ donation and π back-bonding can be separated, and thereby used to derive more complex multi-parameter models. We review here some of the work that has been carried out on such attempts to unravel the relative contribution of the two bonding types in different ligands and different environment, at both a qualitative and quantitative level.

The early suggestions of back-bonding to phosphines related to the very electron-poor PF_3 ligand [81], but it was suggested later, based on the much higher intrinsic stability of *cis* versus *trans* $PtCl_2(PEt_3)_2$ [82], that alkylphosphines should be π acceptors also. As back-bonding to such ligands is rather weak, however, there has been some discussion as to whether simple alkyl and aryl phosphines really do lead to any back-bonding. Correlation between metal–ligand bond energies or related properties, and ligand descriptors that are expected to relate only to the σ donor character, are often quite good. For example, as mentioned above, calculated L–$Ni(CO)_3$ bond energies correlate well with the V_{min} parameter [70], as do the reduction potentials for $[\eta$-Cp(CO)L(COMe)Fe$]^+$ [17]. Many experimental properties also correlate well with σ-only properties for alkyl- and arylphosphines, so it has been suggested that such ligands involve only σ donation [12].

Computation provides unequivocal evidence that back-bonding does occur towards simple trialkyl- and triaryl-phosphines. This is shown by some of the quantitative analyses discussed below, but also by qualitative properties. For example, the calculated electron density in a number of metal–phosphine complexes has been analysed [85]. In cases where back-bonding occurs, a larger density is expected in the region corresponding to overlap between the metal d orbital and the ligand P–C σ* orbitals. Such subtle effects on the total density may be difficult to detect in absolute density plots, as they are quite small, and may be swamped in the overall electron density. However, appropriate plots of electron density *differences* can make such changes apparent.

A particularly useful approach is to plot the density difference between two metal complexes that differ only by the removal of an electron from the metal centre, formally a metal d electron. In the more oxidized complex, either one of the d electrons responsible for back-bonding is removed, or the general lowering of the d orbital manifold due to oxidation of another d orbital leads to decreased back-bonding. In either case, a plot of the density difference shows a main feature corresponding to the particular d orbital from which the electron was removed, and a set of secondary features corresponding to polarization of the ligands, and, where back-bonding was important in the reduced compound, to a loss of density from the region close to the π acceptor ligands of the metal. This is shown schematically in Fig. 7.

An isocontour plot of the electron density difference clearly shows this effect for a number of metal–phosphine complexes. Plots for $CrCl_2(dmpe)_2$, $Mo(CO)_5(PH_3)$ and $Mo(CO)_5(NH_3)$ are shown in the literature [85], with the first two complexes showing significant regions of electron density loss around the expected location of the P–R σ* orbitals. The molybdenum complexes also show regions of density depletion around the CO π* orbitals; these regions are larger than for the phosphine ligands, as expected given the much higher back-bonding accepting character of the carbonyl ligand. The ammonia complex shows no sign of back-bonding to the N–H σ* orbitals. A similar plot is shown below (Fig. 8) for Cp*Fe(dppe), where the density loss in the region of the P–C σ* orbitals is also apparent.

The compounds for which the electron density differences upon oxidation were examined in this study [85] were chosen because the X-ray structure of both the oxidized and reduced forms was known, and it had previously been shown [86] that the structural changes upon oxidation were those that were expected upon experiencing marked decrease in back-bonding interactions. Most notably, oxidation leads to an *increase* in the M–P distance, consistent with a weakening of the bond, whereas for σ-only ligands, a decrease in bond length is observed, due to stronger Coulombic interactions with the more highly charged metal. These geometry changes are reproduced by the calculations [85], so that structural and electronic structure considerations both support a net loss of back-bonding upon oxidation at the metal.

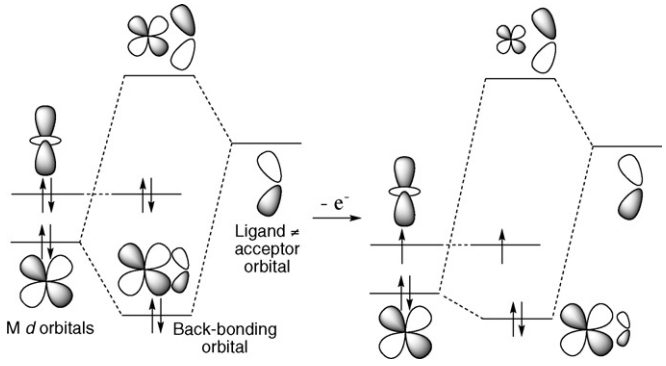


Fig. 7. Schematic molecular orbital diagram showing reduced back-bonding upon one-electron oxidation of a generic metal complex due to lowering of the d orbital manifold.

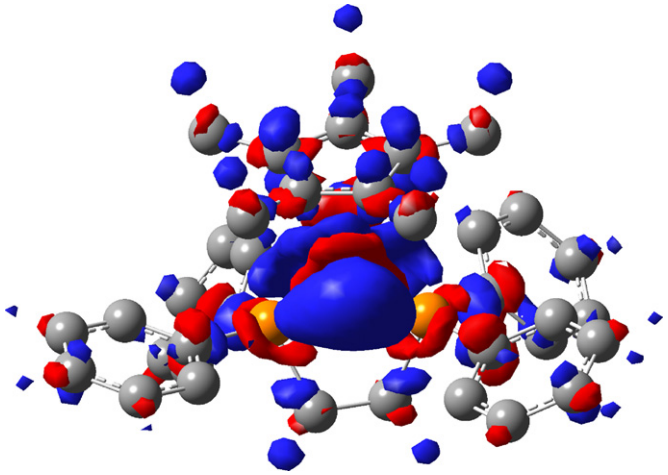


Fig. 8. Density difference plot for removal of an electron from Cp*Fe(dppe), drawn using an isocontour value of 0.002 e $Bohr^{-3}$. Blue regions correspond to regions of electron density depletion, red regions to regions of increase in density. Hydrogen atoms are omitted for clarity.

These electron density and structural effects show that simple alkyl- and aryl-phosphines do bind to transition metals in a complex way involving both σ bonding and π back-bonding. However, such analyses do not provide quantitative insight into the energetic consequences of back-bonding. Other computational techniques are needed to analyse this aspect of bonding.

Computational chemistry is regularly used to calculate the *total* bond energy between a metal fragment and a ligand. Although it is significantly harder to calculate accurate binding energies for transition metal compounds than for many main-group species, computation does provide a very useful estimate of key bond energies. Trends in bond energies for related ligands can in particular often be predicted quite accurately, as error cancellation can be relied upon. A more challenging problem is to calculate the contribution to the bond energy from effects such as back-bonding. The difficulty here is essentially that one is trying to calculate the energy difference between one well-defined, experimentally observable state (the fully bound ligand), and one purely hypothetical state (the complex in which all bonding interactions between the metal and the ligand are present *except* for back-bonding). The latter state cannot in principle be observed and this also leads to serious problems when attempting to calculate its energy.

One approach that has been used by computational chemists to separate bonding into different components is based on comparing the energy of the fully optimised molecular species, and that of a point on the potential energy surface in which the desired interaction is absent due to stereoelectronic effects. For example, lone pair stabilisation of carbocations in α-amino or α-phosphino carbocations $R_2C^+ - XR'_2$ (X = N or P) can be assessed (see *e.g.* [87]) by comparing the energy of the global energy minimum with that of the lowest energy geometry in which the empty p orbital on C and the lone pair on X lie at right angles to each other. This approach seems at first very attractive for metal–phosphine complexes, but is not really applicable due to the high symmetry around the M–P axis, which means that there is no possible angle in which back-bonding interactions are prohibited.

Hence indirect techniques, in which the interaction of interest is isolated from any others as part of the quantum chemical treatment, are needed. Many of these are based on stepwise processes to calculate the energy of interaction between the isolated ligand and the metal fragments, using either Hartree-Fock or, more commonly, DFT methods. First, the ligand and metal fragment are distorted from their equilibrium geometries to the geometry they adopt in the complex. Then, the classical electrostatic interaction energy between the two partners is calculated based on the fragment wavefunctions. Then the two wavefunctions are antisymmetrized, and finally, they are allowed to relax self-consistently to the final function. This relaxation process can be done stepwise, including only parts of the Fock matrix, *e.g.* those allowing mixing between filled ligand orbitals and vacant metal fragment orbitals, and in this manner, a defined energy contribution can in principle be obtained for a particular type of bonding interaction. In another variation, the energy effect of relaxation of orbitals can be broken up into contributions from each of the irreducible representations of the molecular point group of the whole complex, again allowing some distinction to be made between different interactions where symmetry can be used to discriminate between them.

This general description covers methods such as the energy decomposition analysis (EDA) [88], the extended transition state (ETS) model [89], and the constrained space orbital variation (CSOV) method [90]. All of these have been applied to study bonding in transition metal complexes. Another approach, called the Natural Orbitals for Chemical Valence (NOCV) method [91], has also been recently applied to ligand binding. It relies on the calculation of the eigenvectors of the difference density matrix formed by subtracting the density matrices of the fragments from that of the complex. The NOCV method does not provide an energy for a given interaction, but does provide an approximate valency associated with each type of orbital interaction.

The EDA, ETS, and CSOV methods all suggest that bonding of phosphines and related phosphorus ligands to metal complexes involves very significant contributions from back-bonding. For example, back-bonding is predicted using the ETS method to contribute ca. 15 kcal/mol to the bond energy in the three complexes $M(CO)_5(PH_3)$ (M = Cr, Mo and W) [92]. This is a significant contribution, given that the total bond energies for these three complexes, calculated at the same BP86/TZP level of theory, are only just above 30 kcal/mol. The same method predicts similar back-bonding contributions for the PMe_3 complexes, and larger contributions of ca. 30 and 25 kcal/mol, respectively, for the PF_3 and PCl_3 species. In the case of the phosphorus trifluoride compounds, the contribution from back-bonding is roughly equal to the total bond energy, suggesting that in the absence of this interaction, PF_3 would not bind to these metal centres. Clearly, donation of a metal lone pair into the P–X σ^* orbitals is an important contribution to ligand binding. Some caution is however needed when interpreting these numbers, as it has been shown that the ETS predicts a significant contribution of 4 kcal/mol from back-bonding in $M(CO)_5(NH_3)$ (M = Cr, Mo) [93]. Most other experimental evidence, and calculations, including the density difference study reported above, suggest that back-bonding to amine ligands is negligible. Also, the ETS predicts a significant increase in the strength of the back-bonding interaction (to ca 6 kcal/mol) upon ionization to $M(CO)_5(NH_3)^+$ [93]. Structural and density difference plots uniformly predict that back-bonding should decrease upon oxidation.

How can one explain these unexpected results from the ETS method? This approach is based on a decomposition of the total energy difference upon binding into a sum of several terms, corresponding to the individual steps in the hypothetical process described above. Each of these terms is assigned a physical meaning. For example, the energy change corresponding to the orbital relaxation step is usually attributed to covalent bonding, including the back-bonding interaction of interest here. But this identification of the energy difference between two theoretical constructs to a 'chemical' concept such as binding is fraught with difficulties. For example, upon bringing two closed-shell partners such as two helium atoms together, exchange-repulsion due to the antisymmetrization of the two overlapping wavefunctions explains the repulsive nature of the interaction. Polarization of each atom will act to minimise this repulsion. In the ETS formalism, this polarization will appear as an orbital relaxation term, yet most chemists would choose *not* to interpret this as a covalent bonding effect. Likewise, one can imagine that upon complexation of an amine to a metal fragment, a close approach of the nitrogen to the metal will favour a strong dative interaction between the amine lone pair and the empty metal acceptor orbital. This close approach will however be unfavourable in other respects, as it will enforce overlap between orbitals of π symmetry. The associated wavefunction antisymmetrization will lead to a destabilisation term. Here too, this will be alleviated by orbital relaxation—but as for the case where two helium atoms are forced to lie close to each other, it is not appropriate to interpret this effect as true chemical bonding.

The CSOV method, applied using Hartree-Fock calculations, also predicts a significant back-bonding contribution, of 9–20 kcal/mol, in complexes $Pd–PR_3$, where R = H, Me, OMe or F [94]. The CSOV method allows for stepwise mixing of different orbitals, so that it is possible to separate the effect of intra-phosphine relaxation and metal to phosphine back-donation. The relaxation term was found to be fairly small for the phosphine, but fairly large for the metal fragment, which in Ref. [94] was simply a Pd atom.

Finally, the NCOV method has been used to predict the valency corresponding to σ donation and π back-bonding for a set of different ligands in complexes $Ni(NH_3)_3L^{2+}$ and $Ni(CO)_3L^{2+}$ [95]. This method indicates a significant degree of back-bonding to PH_3, with computed bond orders of 0.44 and 0.41, respectively, in the ammine and carbonyl species. As for the ETS approach, the NCOV method does not allow an easy discrimination between intra-ligand and intra-metal fragment relaxation or polarization effects, and actual bonding, so the numerical values of the valences must be considered with some caution. For example, back-bonding to NH_3 is predicted to be almost as strong as to PH_3 [95]. Nevertheless, this approach provides further confirmation of the role of back-bonding in metal complexes.

One other approach that has been used to attempt to quantify the contribution of back-bonding to the metal–phosphorus bond is natural bonding orbital analysis [96,97], and in particular the use of perturbation theory to quantify secondary orbital interactions. In this method, the calculated wavefunction is fitted to a representation in terms of a minimal basis of natural atomic orbitals combined to form molecular bonding orbitals. Then the energetic contribution of deviations from this reference Lewis structure are derived assuming perturbation theory. For transition metal compounds, a good reference representation can be obtained in which pairs of *trans* ligands are treated as bonding to the metal through three-centre four-electron bonds [97]. The t_{2g} orbitals on the metal are treated as lone pairs. Back-bonding then appears as a secondary orbital mixing, and the energy contribution of this mixing gives a good approximation to the energetic impact of back-bonding [93].

This NBO analysis has been applied to a set of model compounds Pd–L, L′–Pd–L, and $M(CO)_5L$, where L is a ligand such as NH_3, PH_3, $P(OMe)_3$ or PF_3, L′ is CO or NH_3, and M is Cr or Mo [93]. The results provide useful insight into the π acceptor character of each ligand, and, by taking the difference with respect to the total bond energy into account, also illuminate the primary σ donor interaction. Perhaps the main observation is that back-bonding occurs with significant strength to all the phosphorus-based ligands, in

agreement with the qualitative conclusion drawn from structural and density factors, and with the quantitative insight derived from ETS and related analyses. A good model of phosphines and related ligands must of necessity take into account their π acceptor character.

The NBO calculations show the amine ligand to lead to weak back-bonding, *e.g.* of 5.7 kcal/mol in the electron rich H_3N–Pd–NH_3 complex, but this is much smaller than the contributions of 16.0 and 14.9 kcal/mol obtained for the H_3N–Pd–PH_3 and Pd–PH_3 complexes, respectively [93]. The total M–PH_3 bond energy in the latter complex is 36.1 kcal/mol, so back-bonding accounts for almost half of the bond energy. In the $Cr(CO)_5L$ complexes, back-bonding makes a much smaller contribution for all the ligands, as they must now compete with the strongly π-accepting CO ligands for the metal d electrons. For ammonia, back-bonding has an almost negligible effect of 0.1 kcal/mol, whereas for PH_3 it accounts for 4.4 of the bond energy, and for PCl_3, it contributes an impressive 15.7 kcal/mol to binding. These results demonstrate that as suggested above for a given ligand, back-bonding can indeed vary in magnitude and in relative contribution to the total bond energy when comparing different metal fragments. Bonding to the Pd atom is much more sensitive to back-bonding than is the metal–phosphorus bond in the $Cr(CO)_5$–L complexes.

Additional interesting trends can be observed. For example, PMe_3 is computed to be a slightly *better* π acceptor than PPh_3 in the Cr and Mo complexes [93], contrary to the expectation based on the fact that PPh_3 has a larger TEP [3,98]. This could be explained if the PPh_3 ligand is a poorer σ donor than PMe_3. Another observation is that back-bonding contributes roughly 50% more to bonding in Pd–PF_3 (23.7 kcal/mol) than in Pd–PH_3 (14.9 kcal/mol), whereas in the $Cr(CO)_5$–L ligands, the difference is much greater, with contributions of 11.9 and 4.4 kcal/mol, respectively. Apparently, compared to PF_3 the weak PH_3 ligand is able to back-bond more effectively in the more electron rich environment of the complex with atomic palladium than in the more electron-poor environment. This effect probably also explains why even ammonia acts as a π acceptor in the Pd complex, although this type of behaviour is not noted for more common amine complexes. Although this conclusion is based on calculations with the somewhat unrealistic model compounds with Pd, it should also apply when comparing 'real' complexes with more and less electron rich metal centres.

As a final note, it is worth highlighting that this section has focused on the way in which computation can provide further insight into the impact of ligand electronic properties on bond energies. This may appear at first sight to be a relatively unimportant property for determining effects such as catalytic efficiency, as it should only matter in the sense that low metal–ligand bond energies can lead to low catalyst stability. In fact, bond energies are much more important than this, as shown in the hypothetical thermodynamic cycle shown in Fig. 9. For any given ligand L, the activation enthalpy $\Delta H^{\ddagger}(L)$ can be written as the activation enthalpy for a hypothetical ligand-free reaction, $\Delta H^{\ddagger}(/)$, plus the difference between the bond energies of the ligand towards the metal in the reactant and in the TS. If the TS is more electron-poor than the reactant, for example, a stronger σ donor character for the ligand may decrease the activation enthalpy, whereas a stronger π acceptor character may increase it.

The more detailed computational investigations of metal–phosphorus interactions described in this section essentially confirm established descriptions of bonding in terms of σ donation and π back-bonding. However, by enabling comparisons between different ligand types and different coordination environments, they suggest that the mix of these different contributions can vary from case to case. The different methods for decomposition of energy contributions are very time-consuming, and in some cases computationally expensive, which restricts the application of such approaches to a few representative ligands and their complexes. Practical explorations of ligand effects for large databases of synthetically relevant ligands and complexes generally rely on descriptors capturing ligand properties implicitly, often by measuring changes in structures, energies and experimentally observable properties, and decomposition into steric and σ/π effects is rarely attempted. The development, analysis and potential applications of such descriptor databases in interpretation, discovery and prediction will be discussed in the next section.

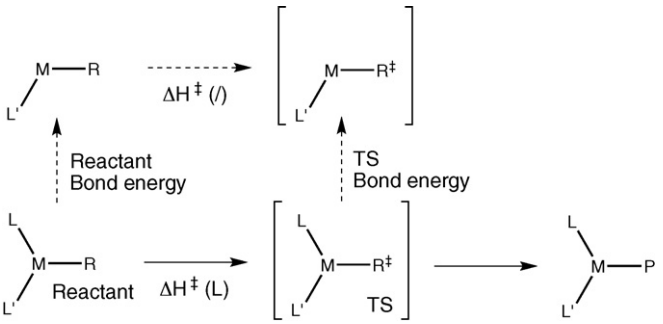


Fig. 9. Notional thermodynamic cycle showing the importance of bond energies for understanding ligand effects on reactivity.

5. Descriptor databases and analysis

The identification and analysis of linear free energy relationships (LFER) has a long tradition in organic chemistry, as they allow the derivation of quantitative relationships between suitable property descriptors and experimental measurements such as rate and equilibrium constants and spectroscopic data [30,34,44]. Similar analyses have been applied to the effects of exchanging P-donor ligands in transition metal chemistry (see for example Refs. [4,25,31,34,35]), using both the individual descriptors summarised above and more extensive databases of ligand descriptors, the latter more likely perhaps to capture the responsiveness of ligand properties to the coordination environment (Section 4). Such databases, along with their applications in data analysis and predictions, will be discussed in this section. While our main focus remains on calculated descriptors, we will also consider parameters derived from experimental data with a view to providing a more complete overview of data analysis techniques. The chosen methods of data analysis and presentation are influenced by the aims of each study, and we will structure this section according to whether interpretation, discovery or prediction is the main focus.

5.1. Interpretation of ligand effects

Where linear relationships between experimental data and suitable descriptors can be identified, the magnitude of contributions from individual descriptors can improve our understanding of ligand effects on the chemical processes (*e.g.* spectroscopic, kinetic, thermodynamic) observed. In Sections 2 and 3 we have explored linear relationships between observations and *individual* descriptors, but, as highlighted in Section 4, chemical reality is often more complicated and several properties, expressed for example in terms of steric and σ and π electronic effects, may contribute. Moving beyond bivariate linear correlations, most analyses aimed at interpreting descriptor contributions use multivariate linear regression [42,43], where a linear equation of the general form

$$P_{obs} = a\,D_1 + b\,D_2 + c\,D_3 + \ldots + \text{constant} \quad (7)$$

where P_{obs} is the observed property and D_1, D_2, D_3 denote different descriptors, is used to model observed data. The quality of fit of this model to the observed data can be assessed by regression coefficients (R^2), where values close to 1 indicate that the model reproduces the observed data well. In addition, if the number of descriptors needed to produce a good model is of interest, for example to avoid overfitting [99], an adjusted version of the regression coefficient (adjusted R^2) can be used to compare different models. This diagnostic takes into account that additional variables will always improve the fit of a model but simpler models are usually more desirable as they are easier to interpret and less prone to perturbation by the effects of noise. While diagnostic plots of fitted versus experimental and residual versus fitted data can be used to illustrate the quality of a multivariate linear regression model, these are of greater interest where predictions are attempted (Section 5.3).

The QALE (quantitative analysis of ligand effects) approach has been applied by Giering and co-workers to experimental data over a period of almost two decades [11–23], during which both descriptors and analyses have been refined; we will concentrate on more recent implementations of this approach here. The QALE model assumes a linear relationship between experimentally measured properties and a number of steric and electronic descriptors, termed stereoelectronic parameters, of the general form:

$$\text{property} = a\,\chi_d + b(\theta - \theta_{st})\lambda + c\,E_{ar} + d\,\pi_p + e\,i + f \quad (8)$$

where χ_d describes σ electron donor capacity [12,13,19], θ is the Tolman cone angle [3], θ_{st} is a steric threshold, λ is a switching function for the steric term, E_{ar} is a secondary electronic effect, originally described as aryl effect [14], π_p describes the π acidity [19], and i describes the number of hydrogens in $PZ_{3-i}H_i$ [23]. Steric effects can be discontinuous, only becoming important after ligands have reached a critical size (a *steric threshold*). While the coefficients *a–f* in the QALE equation were initially determined from a combination of graphical and regression analyses for different types of ligands (summarised in Ref. [15]), later applications relied on fitting appropriate regression models for multiple variables [21]. The stereoelectronic parameters χ_d, θ, E_{ar} and π_p have been reported for 138 P-donor ligands in the literature [20] and for over 300 ligands on the QALE website (http://www.bu.edu/qale/) [23]. Some of these parameters were derived from experimental measurements [3,14,46,53] and others were estimated by reversing the QALE analysis and predicting parameter values from fitted models [18–20].

Analyses based on the QALE equation (Eq. (8)) have been useful in formalising the interpretation of many experimental datasets in terms of steric and electronic ligand effects, and much of the recent work published by the authors [18–20] routinely examined between 15 and 32 experimental (spectroscopic, electrochemical and thermochemical) datasets for a variety of phosphorus donor ligands. A recent analysis of π effects in rhodium complexes used a more targeted set of data, considering spectroscopic (ν_{CO}), structural (Rh-P distances) and thermodynamic ($-\Delta H_{rx}$) observations for $[Rh(CO)Cl(PA_3)_2]$ and $[Rh(acac)(CO)(PA_3)]$ complexes [21]. Regression coefficients R^2 between 0.961 and 0.995 were achieved for the models; the carbonyl stretching frequencies ν_{CO} gave the best models ($R^2 = 0.990$ and 0.995) with QALE parameters. The model coefficients were then standardised by considering the range of coefficients determined for a large set of models, and the percent contribution of descriptors to the models could thus be analysed. These contributions suggest that π back-bonding is important in rhodium complexes with phosphorus donor ligands. However, the balance of effects due to different co-ligands complicates the analysis such that π effects can be observed for all datasets apart from $-\Delta H_{rx}$ for the ligand exchange $[RhCl(CO)_2]_2 + 4PA_3 \rightarrow 2[RhCl(CO)(PA_3)_2] + 2CO$, where the CO ligands may balance the effects of changes in Rh-P π back-bonding; such effects appear more clearly for complexes with acac ligands; again, ligands appear to respond to changes in the coordination environment.

The QALE approach has been applied by other groups as well, most notably by Poë and co-workers in their analysis of associative reactions of metal carbonyl clusters [24-30]. These authors used a slightly different σ electronic parameter pK'_a, linked more closely to experimental pK_a values [25,28,46] but considering Tolman's descriptors [3,53] in the derivation, as well as a different π parameter $pK'_a\pi$ [26] and an additional term for pendant groups attached to the ligand, ΩE_{pge} [29,30] to give:

$$\text{property} = \alpha + \beta(pK'_a + 4) + \gamma(\theta - \theta_{th})\lambda + \phi(pK'_a\pi) + \delta(E_{ar}) + \Omega E_{pge} \quad (9)$$

They also calculated the contribution of each parameter to an experimentally observed property for individual ligands [27] rather than using ranges to standardise data [21], but otherwise they followed the QALE protocol. Recently, Bunten and Poë have published a discussion of multi-mechanism linear free energy relationships (MMLFER) for the analysis of reduction potential data for the $[(\eta^5\text{-}C_5H_5)(CO)(L)Fe(COMe)]^{+/0}$ couple with a range of 22 phosphorus donor ligands L [30], which compared the two flavours of QALE in an appendix.

Another system of descriptors for linear free energy relationships was applied to P-donor ligands by Drago and Joerg [32,33]. These descriptors have been derived from the enthalpy of adduct formation with different fragments and seek to capture the σ donor strengths of ligands in two parameters, E_B and C_B, designed to measure electrostatic and covalent character, respectively, which have been reported for 37 P-donor ligands. The so-called ECW model, named after terms in the model equation

$$\text{property} = E_A E_B + C_A C_B + W \quad (10)$$

where subscripts A denote acids and B denote bases was originally developed for many different classes of ligands, with the inclusion of P-donor ligands a later development requiring different descriptor derivations [32,33]. Models were derived for a wide range of experimental datasets, including heats of reaction, ^{13}C NMR chemical shifts, carbonyl stretching frequencies, redox potential and rate constants for ligand displacements. This system was presented as a simpler, chemically more intuitive way of analysing experimental data for a variety of ligands, in direct opposition to QALE and offering alternative interpretations of experimental data in terms of ligand donor/acceptor strength in a ligand set extending beyond P-donors [32]. This precipitated a debate in the literature [16,17,33], during which Giering and co-workers claimed that two parameters may not be sufficient to capture the steric and electronic properties of P-donor ligands [16,17] and Drago and Joerg questioned the consistency of interpretations derived from QALE analyses [33]. While other authors have criticised the QALE approach for the derivation of ligand subsets and the treatment of outliers [31], as well as the complicated analysis required [38], QALE has received more widespread attention in the analysis of P-donor ligand effects on experimental data than the ECW model and has been described as a superior approach [26,38], helped perhaps by the fact that descriptors are more clearly targeted at P-donor ligands.

Brown combined his calculated steric descriptor E_R (described above [34,54,63–65]) with experimental ^{13}C NMR chemical shifts of the carbonyl group in $[Ni(CO)_3L]$ complexes, δ [55], to derive correlation coefficients and linear free energy relationships for just two descriptors [34,54], mainly considering rate constants for ligand

exchange reactions, *e.g.* $[Co(NO)(CO)_3] + L \rightarrow [Co(NO)(CO)_2L] + CO$ [45] and $[Fe(NO)_2(CO)_2] + L \rightarrow [Fe(NO)_2(CO)L] + CO$ [100]. The analyses did allow for steric thresholds, but he did not attempt to separate electronic contributions into σ and π effects, making a simpler two descriptor model more acceptable (see, however, Ref. [16]). The main aim of this work was to establish and promote E_R as a useful steric descriptor and little discussion of the descriptor contributions and models was included. However, to allow comparison of different reactions Brown proposed a standard set of 10 ligands, consisting of phosphines and phosphites with different alkyl and aryl substituents and hence capturing a range of properties, for future experimental explorations of ligand effects [34].

Any attempt to analyse the contribution of ligand descriptors to experimental observables relies on the availability of suitable datasets. Such data should have been determined consistently, subject to the same experimental conditions, and ideally a large number of ligands would have been sampled to allow for meaningful statistical analysis. These criteria are not always fulfilled by available experimental datasets, which can make their analysis difficult. In addition, if different subsets of ligands are considered in different experimental datasets, this will also affect the descriptor coefficients in fitted equations. While *e.g.* QALE analysis tries to take this into account by standardising each coefficient and comparing percent contributions [21], these coefficients are unlikely to be particularly robust with respect to changes in the ligand set and such comparisons can only be qualitative, unless the same ligand set has been considered throughout; this could be achieved with a standard ligand set as suggested by Brown and Lee [34]. Calculating parameter coefficients for individual ligands and considering them as modifiers of a standard reaction rate as described by the Poë group [27] may be a more suitable approach, although the generality of this approach was not explored. Furthermore, the experimental sampling of ligands is often not evenly distributed between alkyl- and arylphosphines, phosphites, aminophosphine and halogenated phosphines, so ligand effects important for a subset may be overemphasised in an analysis, if only few examples of other ligand types have been considered. Indeed, Drago and Joerg make a similar point with respect to wider ligand space, *i.e.* combining ligands with different donor atoms (phosphorus, oxygen, nitrogen, sulphur) to analyse acceptor parameters more reliably [32,33], although in practice few of the experimental datasets they considered actually sampled more than one subset of ligand types comprehensively.

5.2. Discovery of ligands

While the linear free energy relationships discussed above often seek to unravel the contributions of ligand steric and electronic properties to experimental observations, descriptors can also be used to illustrate ligand similarities. Moving beyond inspection of individual descriptor values, Tolman's stereoelectronic map of P-donor ligands (Fig. 2) arises from a straightforward scatter plot of his two descriptors, cone angle θ and the electronic parameter ν/TEP. The resulting map of ligand space can be used to locate ligands with similar properties to a ligand of interest or to inspire a range of experiments with very different ligands, both concepts useful in experimental design, optimisation and ligand screening. For such maps, a broad and even scatter of data points is desirable, as this indicates that the ligands sampled cover a range of properties and that descriptors are not correlated; high linear correlation and regression coefficients would be undesirable. Where multiple descriptors have been determined, two- or three-dimensional scatter plots of descriptor combinations may not capture sufficient information and become unwieldy, so statistical projection methods such as principal component analysis (PCA) [42,43] can be more useful. PCA generates derived variables, principal components (PCs), as linear combinations of the original descriptors, subject to capturing most of the variation in the original dataset in the first few PCs. In addition, the PCs are uncorrelated, making them ideal for reducing the dimensionality of a large set of descriptors. Scatter plots of the first two or three PCs often capture a large proportion of the variation and hence information content of underlying descriptors. However, a simple and intuitive interpretation of axes in terms of familiar steric and electronic effects becomes difficult for such derived descriptors and PCA is known to be sensitive to outlier observations [74,75].

Principal component analysis was used in work by Bjørsvik et al. to design a test set of ligands for more detailed predictive analysis (discussed below) [101]. They calculated a range of structural and electronic descriptors (*e.g.* FMO energies, hardness, dipole moment, number of C and H atoms, surface area, *etc.*) with semi-empirical methods for 45 P-donor ligands and then used PCA to derive an initial map of ligands, which showed two distinct clusters. This led to the exclusion of the smaller cluster of polyhalogenated ligands to give a revised ligand map of 39 ligands, from which a subset of 17 ligands was chosen as training set for predictive regression analysis (Section 5.3). While in this case the PCA map was secondary to the predictive models developed, this work demonstrated how such maps can be used in experimental design to ensure that a range of properties is captured by the ligand set investigated.

Cundari and co-workers were interested in the computational evaluation of new ligands and they used the computational steric and electronic descriptors S4′ and SEP described above to generate a stereoelectronic map of P-donor ligand space [68] (Fig. 10a). On this map, they identified areas not sampled well by com-

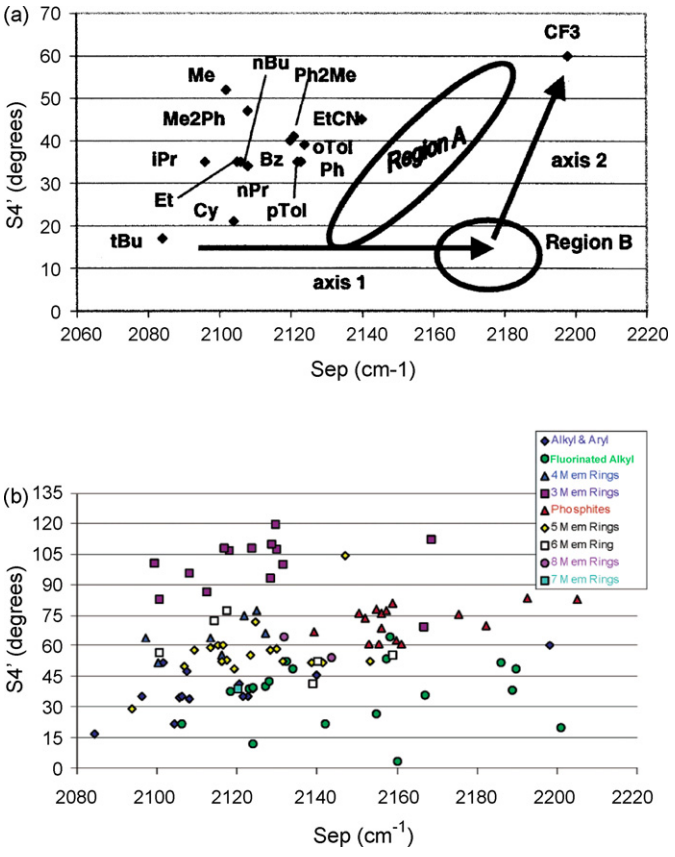


Fig. 10. Stereoelectronic maps based on SEP and S4′: (a) map used in design and (b) expanded map. Reprinted with permission of the copyright holders from Ref. [68]. Copyright 2003 American Chemical Society.

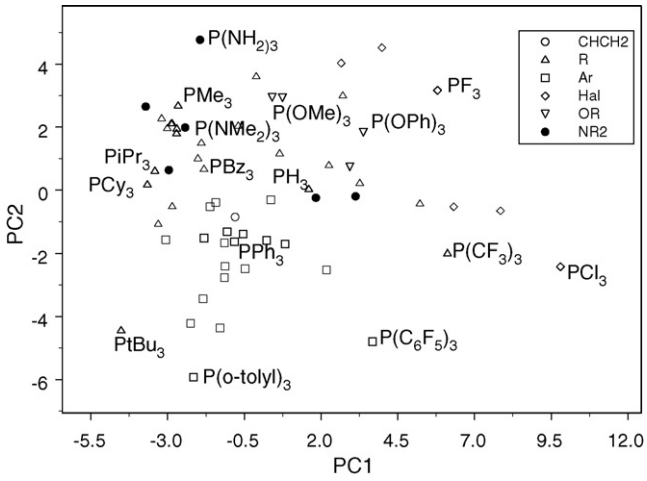


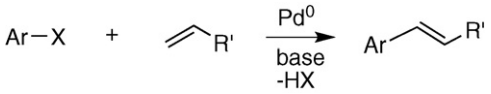
Fig. 11. Ligand knowledge base map derived from principal component analysis on 23 descriptors; see Ref. [74] for map incl. Varimax rotation and further details of ligands and analyses.

monly used ligands and translated this into design criteria for phosphines and phosphites, targeting especially large, electron-poor phosphines and ligands of median electronic and donor properties. By exploring the experimental literature and CSD [73] for less popular ligands, they then identified a set of target ligands (phosphites, perfluorinated phosphines and ligands where the phosphorus is part of a ring) and calculated descriptors for these ligands, producing an extended stereoelectronic map (Fig. 10b). This allowed them to identify several candidate ligands for further experimental investigation, including P(t-C_4F_9)$_3$ as a large, electron-poor ligand, as well as a range of cyclic phosphines.

To exploit the information content of the calculated descriptors in a ligand knowledge base context (see above and Refs. [74,75]), we have used principal component analysis on the descriptor set shown in Table 3 [74]. The original 23 descriptors, derived from both the free ligands and a range of metal and Lewis acid fragments (Table 3) are correlated to some extent, whereas PCA removes this correlation and captures 67% of the variation in the dataset in the first two PCs. The resulting map of ligand space (Fig. 11) shows chemically intuitive clustering of ligands according to their substituents and can be used to identify ligand similarities and design test sets for screening and reaction optimisation.

Using steric (S_{eff}) and electronic (E_{eff}) descriptors derived from V_{min} calculations with a combination of QM and MM approaches, Suresh and co-workers have also produced a map of phosphine ligands (Fig. 12) [72]. As in Cundari's map [68] described above (Fig. 10), they were interested in locating unusual ligands retrieved from the CSD in relation to more familiar ligand families. They concentrated on cyclic and heterocyclic phosphines and phosphites, as well as some novel ligand structures reported recently in the literature which they related to more familiar ligands. While the main aim of this work was to introduce and describe their V_{min} derived descriptors, they briefly discussed the use of such a map in catalyst design and selection.

Work by Rothenberg and co-workers has explored the application of calculated QSAR descriptors and multivariate analyses in transition metal catalyst screening [102–104]. They calculated 80 structural and electronic descriptors for P-donor ligands and solvents and used these to produce PCA maps and PLS models for palladium-catalysed Heck reactions (Scheme 2 [105]). They demonstrated how a small training set of 9 ligands could be used to screen other ligands based on a ligand map, where they projected a virtual



Scheme 2. Heck reaction.

library of ligands to help identify promising candidates for further screening. They also discussed how a PCA-derived solvent chart might be used to optimise reactions. However, their work is mainly targeted at making predictions and will be discussed in detail in Section 5.3.

These studies have demonstrated how ligand properties can be visualised in scatter plots of steric and electronic descriptors, or of projected variables when the underlying database is multidimensional. Such maps of ligand space are descendants of Tolman's stereoelectronic map (Fig. 2 [3]) but better access to calculated descriptors and the possibility of using more dimensions to describe ligand space and its properties has facilitated the characterization of novel or unusual ligands by relating them to more familiar species. The underlying assumption here is that spatial proximity on a ligand map arises from similar steric and electronic properties and that such maps can guide experimental efforts, either to focus on areas of ligand space which contain ligands useful to a specific application, *e.g.* in catalysis, or to explore regions not considered previously. As indicated in Section 4, ligand properties may not be captured reliably by just two descriptors, as they can be responsive to changes in their coordination environment, and we would expect projections of larger descriptor sets to be more meaningful in that respect. Successfully translating such maps into predictions (of activity, selectivity, properties, *etc.*) is the logical next step in applying calculated ligand descriptors, and this will be reviewed in the next section.

5.3. Prediction of ligand effects

The determination of quantitative structure–activity relationships (QSAR) and related structure–property (QSPR) and structure–selectivity (QSSR) relationships has been a key tool in medicinal chemistry, where large databases of compounds may need to be evaluated for their biological activity in preparation for high-throughput screening [106]. Where prediction, rather than interpretation and discovery, is the main aim of studying ligand effects, both ligand descriptors and analysis methods may be chosen according to different criteria. More descriptors are usually considered, with less emphasis on their detailed interpretation in terms of familiar steric and σ/π electronic effects, both because such interpretation becomes more complicated as the diversity of ligand sets increases, and because additional descriptors are perhaps more likely to capture ligand properties in different environments, thus giving rise to better models. Often, these descriptors have been calculated with very cheap computational approaches to give access to large databases, and they may be based on topological indices, *e.g.* assessing the number and size of rings in a substrate, rather than considering the three-dimensional structure of molecules (see Refs. [104,106,107] and references cited for overviews and applications of such descriptors). Data analysis often relies on multivariate regression and projection methods, and for the resulting models both data fit and quality of predictions need to be assessed. While multiple linear regression (MLR) models as described above may still be useful, non-linear approaches (artifical neural networks (ANN) and genetic algorithms (GA)) as well as regression models using derived variables (principal component regression (PCR) and partial least squares (PLS)) are now used routinely in chemical data analysis [43,104,106]. For the simpler linear

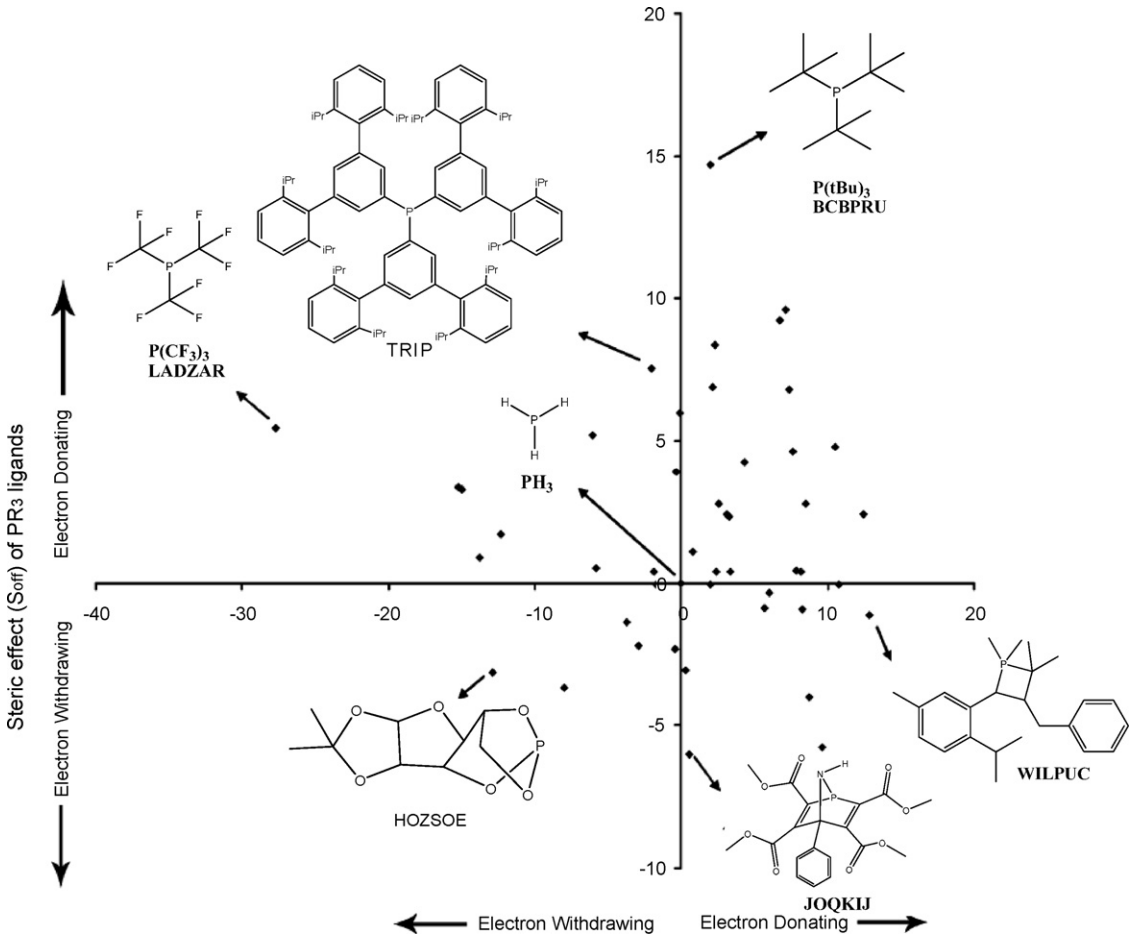


Fig. 12. Stereoelectronic map based on E_{eff} and S_{eff} parameters. Reprinted with permission of the copyright holders from Ref. [72]. Copyright 2007 American Chemical Society.

regression models, correlation between descriptors can give rise to a range of models of comparably good fit to the training data, and model complexity and prediction errors need to be considered to choose a suitable model [75]. Such prediction errors can be determined by fitting models to a subset of the data and assessing predictions for the results not used in the fitting; several iterations of this with subsets of the data may be used if the overall dataset is small and resampling approaches such as cross-validation or the bootstrap have to be used [75]. Any such analysis depends on the availability of suitable external data, sampling a large and varied set of compounds. The application of predictive analysis to organometallic species and more specifically ligand effects is thus still very much in its infancy, with most studies seeking to establish the chosen approach and reporting only limited testing of the resulting predictions.

While many descriptors used for QSAR studies are two-dimensional and topological, three-dimensional descriptors have also been developed to take the molecular geometry into account. Comparative molecular field analysis (CoMFA) [108] is one such 3D QSAR approach which has been applied to P-donor ligands [109]. Starting from a suitable molecular geometry aligned in a standard orientation, the interaction energy between a probe molecule or atom and the structure of interest is determined for a three-dimensional grid surrounding the molecule [110]. Depending on the probe, this grid captures the steric or electrostatic molecular field and the interaction energy at each point can be used as a descriptor in a multivariate regression model, usually partial least squares (PLS). Steinmetz used a standard CoMFA approach to calculate steric and electrostatic fields for a set of 36 P-donor ligands with a variety of substituents [109]. He then derived PLS models for 6 experimental data sets (thermodynamic and kinetic; ligand exchange reactions for 6 different metal centres), which spanned between 8 and 19 ligands and compared the performance of these models with regression models based on cone angles θ and the ^{13}C NMR descriptor δ tabulated by Brown and Lee (see above) [34]. Models were assessed mainly by their squared correlation coefficient, R^2, and the squared cross-validated correlation coefficient, Q^2. The CoMFA models gave acceptable model fit with R^2 between 0.748 and 0.957, but poorer predictive performance (Q^2 between −0.044 and 0.644) than models based on θ and δ (R^2 0.706–0.944, Q^2 0.345–0.896). This was related to the small size of the experimental data sets and poor sampling of parameter space. To establish the validity of the approach, Steinmetz correlated the CoMFA field with these descriptors and Brown's calculated E_R [34,54] (R^2 = 0.924–0.953, Q^2 = 0.681–0.768). While this work demonstrated that CoMFA can be applied to inorganic systems and established a relationship between molecular fields and familiar ligand descriptors, the small size of datasets hampered a more detailed assessment and models gave quite poor predictive performance when compared by cross-validation error estimates.

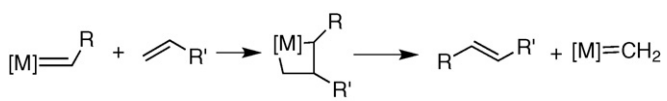
A mixture of QSAR and semi-empirical calculations were used to generate descriptors for another exploratory study of multivariate analyses applied to 45 P-donor ligands, in this case aimed at experimental design [101]. These descriptors only considered the free ligands and were initially assessed with principal

component analysis (see Section 5.2 above) to design a training set of ligands which sampled properties widely and evenly. PLS regression models using two latent variables were then derived for this training set with two experimentally determined IR frequencies (of A_1 and E symmetry for $Ni(CO)_3L$ in CH_2Cl_2) as the responses (R^2 = 0.890 and 0.885, respectively). The models were evaluated by cross-validation (root mean square error of prediction, RMSEP = 3.09 and 4.18 cm^{-1} for ν_{A1} and ν_E, respectively) and prediction of a test set of 12 ligands (RMSEP = 3.60, 5.09 cm^{-1}). Based on these models, predictions were made for all 39 ligands and descriptor contributions were analysed. The authors also discussed the link between their descriptors and familiar ligand properties such as electron donation/back-donation and steric effects.

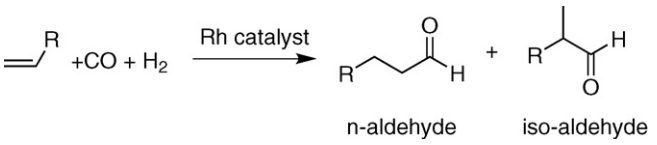
Drawing on 708 traditional and semi-empirical QSPR descriptors, Bosque and Sales explored the use of different regression models for the prediction of ^{31}P NMR chemical shifts for 291 P-donor ligands with alkyl and aryl substituents [111]. Based on extensive model evaluation and testing, they determined that models with seven descriptors would be ideal and used linear regression, neural networks and genetic algorithms to explore different models. These models were evaluated by their regression coefficients, cross-validation error estimates and predictions for an external test set of 30 phosphines. The most advanced neural network/genetic algorithm approach gave the best predictive performance (R^2 = 0.951, R^2(CV) = 0.900, R^2(test) = 0.953, training set error = 9.5 ppm, test set error = 11.3 ppm), suggesting that the NMR chemical shifts can indeed be predicted from descriptors of the molecular structure. As achieving such good prediction was the main focus of this work, descriptor interpretations were not discussed in detail.

As described above, Rothenberg and co-workers have calculated 80 structural and electronic descriptors for P-donor ligands and solvents and used these in maps (Section 5.2) as well as PLS models with turnover numbers (TON) and turnover frequencies (TOF) extracted from the literature for palladium-catalysed Heck reactions (Scheme 2 [105]) as the response data. The 500 reactions they considered were carried out under different conditions and they observed clustering according to such conditions, which prevented them from optimising ligands and solvents within the same model. In a later publication, they applied a similar approach to a set of 412 Heck reactions [103]. They used neural networks to derive the models and analysed descriptor contributions in greater detail. They then applied the models to screen a virtual library of 60,000 reactions derived from 61 new ligands combined with different substrates, catalyst precursors and reaction conditions. They found that non-linear regression techniques gave better models and presented some mechanistic interpretations of the descriptor contributions, relating the importance of palladium loading in the model to the formation of inactive palladium clusters.

A QSAR model has been derived to explore ligand effects in ruthenium catalysed alkene metathesis (Scheme 3 [112]), using calculated catalyst productivity as the response variable [113]. While the main focus was on *N*-heterocyclic carbene ligands, 16 P-donor ligands were also considered. The response was determined by calculating the energy differences between key species along the reaction coordinate with DFT for a representative subset of ligands, including 6 P-donor ligands. The model used a combination of standard QSAR-type topological descriptors and DFT-calculated descriptors relevant to the ruthenium catalysts with partial least squares regression, and good model performance with errors around 1 kcal mol^{-1} for a four variable model is reported (RMSE(CV) = 0.77 kcal mol^{-1}, RMSEP = 0.53 kcal mol^{-1}). The authors attributed the success of their approach to the inclusion of highly reaction-specific descriptors. They have successfully related some of their descriptors to the mechanism of reaction and the stabilisation of the metallacyclobutane intermediate, and hence suggested a number of carbene ligands for further development. Recently Jensen has presented more detailed studies of the mechanism of rhodium-catalysed hydroformylation (Scheme 4) for 11 ligands, including 7 P-donor ligands [114]. These results have also been used in structure–activity relationships aimed at separating steric and electronic effects by using Tolman's descriptors (cone angle θ and TEP/ν) for the P-donor ligands to derive linear regression models for the reaction barrier of key steps. In addition, reaction barriers were calculated for three novel ligands, including $P(OCF_3)_3$, to assess their effect on the reaction pathway and hence their suitability.



Scheme 3. Alkene metathesis.



Scheme 4. Hydroformylation.

As discussed previously, the descriptors in our ligand knowledge base LKB-P (Table 3) [74,75] can be used in a variety of ways, exploring correlations between individual descriptors and other ligand parameters or experimental data (Section 3), deriving maps of ligand space (Section 5.2) or developing predictive models with a variety of approaches (this section). When designing the descriptors for this knowledge base [74], we wanted to achieve robust and reliable calculations to facilitate future expansions and automation (*computational* robustness), as well as covering a range of relevant and varied chemical environments to ensure that descriptors and results are transferable (*chemical* robustness). In addition, we have sought to achieve *statistical* robustness by using approaches suitable for uneven sampling and small datasets [74,75]. Similar considerations have shaped the design of a knowledge base for chelating ligands [76]. However, suitable experimental data remains limited and sampling of chemical space is often uneven, so in the testing of this approach we have favoured simple linear and robust regression models with the aim of developing robust predictive models which remain chemically intuitive [75]. By capturing ligands in a range of chemical environments, steric and electronic properties are implicit in the LKB-P descriptors and we have demonstrated the transferability of these parameters by deriving a range of models for structural, spectroscopic and thermodynamic response data [74], including a more detailed comparison of statistical models for the TEP [75]. These models gave very high regression coefficients (R^2 = 0.988–0.998) and good prediction error estimates through cross-validation and bootstrap methods [74], confirming the chemical robustness of the descriptors. The varied range of applications discussed here and in the preceding sections illustrate that our initial aims in the design of this knowledge base have been met and that a range of descriptors derived from relevant and varied coordination environments is most likely to capture the ability of some ligands to respond by changing their steric and electronic properties.

6. Conclusions

This review has sought to summarise the current state of calculated ligand descriptors for phosphorus donor ligands, including our understanding of the electronic contributions to metal–phosphorus bonding derived from computational studies. Moving the quantitative assessment of ligand properties away from "experimental" parameters such as those described by Tolman [3] towards calculated descriptors opens up new applications in the design and discovery of novel ligands as well as suggesting a new "take" on known ligands by contextualising their properties on ligand maps. A number of examples of such applications have been reported in the literature recently and are reviewed here, and the cost/gain balance for large-scale computational studies continues to improve in favour of such developments.

It is thus tempting to speculate on the extension of such approaches to other ligand types and indeed some of the work discussed here also covers other ligand classes (see for example [8,32,33,36,54,65,77,79,93,107,113–115]) and work on *N*-heterocyclic carbene (NHC) ligands has been reviewed [8,10,41]. For many calculated descriptors, such an extension is reasonably straightforward and simply involves additional calculations with other ligand classes, *e.g.* TEP-related descriptors can be calculated for many other neutral and indeed charged ligands [36,79]. Other descriptors may be less transferable, *e.g.* it has been suggested that a spherical measure of sterics does not capture the asymmetric distribution of steric bulk encountered in NHC ligands well [8,10] and our own work on developing descriptors for P,P- and P,N-donor chelating ligands (LKB-PP) highlighted that Pt(0)-N bonding is too weak to derive suitable descriptors from $[(PH_3)_2Pt\{LL\}]$ complexes [76] analogous to our monodentate LKB-P [74]. Even when the calculations are feasible, inherent differences in bond strengths and lengths may need to be taken into account by calculating descriptors relative to a reference ligand of each class rather than absolute values [76]; otherwise projection methods (*e.g.* PCA) and regression models may be dominated by such changes in bonding and fail to capture relevant ligand properties where ligands are spectators. Nevertheless, such issues can and will be overcome, and we look forward to future developments in this area, giving rise to maps and models for multiple ligand classes and hence extending the scope for ligand design and discovery.

However, the limited availability of suitable experimental data for testing and indeed challenging such maps and models remains a considerable handicap, especially where experimental sampling of ligands is limited to a subset of popular alkyl- and aryl-phosphine ligands, with little consideration given to more exotic ligand architectures, perhalogenated ligands and mixed substitution patterns with OR and NR_2 substituents, nor indeed other ligand classes such as NHC ligands. Much of the work reviewed here hints at the importance of even and extensive sampling of ligand space, both in achieving more reliable and transferable models and predictions, and in order to open up new opportunities in ligand design and discovery. We hope that an overview of the available computational descriptors will help in realising this potential.

Acknowledgements

We would like to thank our co-workers, collaborators and students, M.F. Haddow, A.J. Hamilton, S.E. Harris, D. Hollinshead, D. Hose, D. Lathbury, T. Leyssens, G.C. Lloyd-Jones, R.A. Manson, C.L. McMullin, P. Murray, R. Osborne, G.J.J. Owen-Smith, D. Peeters, M. Purdie, A.C. Tsipis, A. Comas-Vives and A. Welsh, for their contributions to the ligand knowledge base projects and many fruitful discussions. NF and JNH thank the EPSRC for the award of Advanced Research Fellowships (EP/E059376/1 and GR/S51059/01).

References

[1] A.H. Hoveyda, A.W. Hird, M.A. Kacprzynski, Chem. Commun. (2004) 1779.
[2] J.G. de Vries, L. Lefort, Chem. Eur. J. 12 (2006) 4722.
[3] C.A. Tolman, Chem. Rev. 77 (1977) 313.
[4] P.D. Dias, M.E.M. de Piedade, J.A. Marinho Simões, Coord. Chem. Rev. 135/136 (1994) 737.
[5] P.C.J. Kamer, P.W.N.M. van Leeuwen, J.N.H. Reek, Acc. Chem. Res. 34 (2001) 895.
[6] W.A. Herrmann, K. Öfele, D. von Preysing, S.K. Schneider, J. Organomet. Chem. 687 (2003) 229;
R.G. Arrayás, J. Adrio, J.C. Carretero, Angew. Chem. Int. Ed. 45 (2006) 7674;
E.A.B. Kantchev, C.J. O'Brien, M.G. Organ, Angew. Chem. Int. Ed. 46 (2007) 2768;
N. Marion, S. Diez-Gonzalez, S.P. Nolan, Angew. Chem. Int. Ed. 46 (2007) 2988.
[7] B. Schlummer, U. Scholz, Adv. Synth. Catal. 346 (2004) 1599;
U. Christmann, R. Vilar, Angew. Chem. Int. Ed. 44 (2005) 366;
J.-P. Corbet, G. Mignani, Chem. Rev. 106 (2006) 2651.
[8] L. Cavallo, A. Correa, C. Costabile, H. Jacobsen, J. Organomet. Chem. 690 (2005) 5407.
[9] R.H. Crabtree, J. Organomet. Chem. 690 (2005) 5451.
[10] S. Diez-Gonzalez, S.P. Nolan, Coord. Chem. Rev. 251 (2007) 874.
[11] M.N. Golovin, M.M. Rahman, J.E. Belmonte, W.P. Giering, Organometallics 4 (1985) 1981;
M.M. Rhaman, H.Y. Liu, A. Prock, W.P. Giering, Organometallics 6 (1987) 650;
K. Eriks, W.P. Giering, H.-Y. Liu, A. Prock, Inorg. Chem. (1989) 28.
[12] M.M. Rahman, H.-Y. Liu, K. Eriks, A. Prock, W.P. Giering, Organometallics 8 (1989) 1.
[13] H.-Y. Liu, K. Eriks, A. Prock, W.P. Giering, Organometallics 9 (1990) 1758.
[14] M.R. Wilson, D.C. Woska, A. Prock, W.P. Giering, Organometallics 12 (1993) 1742.
[15] J. Bartholomew, A.L. Fernandez, B.A. Lorsbach, M.R. Wilson, A. Prock, W.P. Giering, Organometallics 15 (1996) 295.
[16] A.L. Fernandez, C. Reyes, M.R. Wilson, D. Woska, A. Prock, W.P. Giering, Organometallics 16 (1997) 342.
[17] A.L. Fernandez, T.Y. Lee, C. Reyes, A. Prock, W.P. Giering, Organometallics 17 (1998) 3169.
[18] A.L. Fernandez, T.Y. Lee, C. Reyes, A. Prock, W.P. Giering, C.M. Haar, S.P. Nolan, J. Chem. Soc. Perkin Trans. 2 (1999) 2631;
A.L. Fernandez, C. Reyes, T.Y. Lee, A. Prock, W.P. Giering, C.M. Haar, S.P. Nolan, J. Chem. Soc. Perkin Trans. 2 (2000) 1349;
D. Woska, A. Prock, W.P. Giering, Organometallics 19 (2000) 4629.
[19] A.L. Fernandez, C. Reyes, A. Prock, W.P. Giering, J. Chem. Soc. Perkin Trans. 2 (2000) 1033.
[20] A.L. Fernandez, M.R. Wilson, A. Prock, W.P. Giering, Organometallics 20 (2001) 3429.
[21] M.R. Wilson, A. Prock, W.P. Giering, A.L. Fernandez, C.M. Haar, S.P. Nolan, B.M. Foxman, Organometallics 21 (2002) 2758.
[22] W.P. Giering, A. Prock, A.L. Fernandez, Inorg. Chem. 42 (2003) 8033.
[23] A.L. Fernandez, A. Prock, W.P. Giering, The QALE Web Site, http://www.bu.edu/qale/, last updated 12/09/2003.
[24] D.H. Farrar, A.J. Poë, Y. Zheng, J. Am. Chem. Soc. 116 (1994) 6252;
N.M.J. Brodie, A.J. Poë, Can. J. Chem. 73 (1995) 1187;
R.H.E. Hudson, A.J. Poë, Organometallics 14 (1995) 3238;
A.J. Poë, Y. Zheng, Inorg. Chim. Acta 252 (1996) 311;
R.H.E. Hudson, A.J. Poë, Inorg. Chim. Acta 259 (1997) 257;
C. Babij, H.W. Chen, L.Z. Chen, A.J. Poë, Dalton Trans. (2003) 3184;
C. Babij, L.Z. Chen, I.O. Koshevoy, A.J. Poë, Dalton Trans. (2004) 833;
K.A. Bunten, C. Moreno, A.J. Poë, Dalton Trans. (2005) 1416.
[25] L. Chen, A.J. Poë, Coord. Chem. Rev. 143 (1995) 265.
[26] K.A. Bunten, L. Chen, A.L. Fernandez, A.J. Poë, Coord. Chem. Rev. 233/234 (2002) 41.
[27] K.A. Bunten, D.H. Farrar, A.J. Poë, Organometallics 22 (2003) 3448.
[28] C. Babij, A.J. Poë, J. Phys. Org. Chem. 17 (2004) 162.
[29] K.A. Bunten, A.J. Poë, T.A. Stromnova, Dalton Trans. (2005) 3780.
[30] K.A. Bunten, A.J. Poë, New J. Chem. 30 (2006) 1638.
[31] S.Q. Song, E.C. Alyea, Comm. Inorg. Chem. 18 (1996) 145.
[32] R.S. Drago, S. Joerg, J. Am. Chem. Soc. 118 (1996) 2654, and references cited.
[33] S. Joerg, R.S. Drago, J. Sales, Organometallics 17 (1998) 589, and references cited.
[34] T.L. Brown, K.J. Lee, Coord. Chem. Rev. 128 (1993) 89.
[35] R. Romeo, M.R. Plutino, L.M. Scolaro, S. Stoccoro, Inorg. Chim. Acta 265 (1997) 225.
[36] L. Perrin, E. Clot, O. Eisenstein, J. Loch, R.H. Crabtree, Inorg. Chem. 40 (2001) 5806.
[37] D. White, N.J. Coville, Advances in Organometallic Chemistry, vol. 36, 1994, p. 95.
[38] O. Kühl, Coord. Chem. Rev. 249 (2005) 693.
[39] P. Dierkes, P.W.N.M. van Leeuwen, J. Chem. Soc. Dalton Trans. (1999) 1519;
P.W.N.M. van Leeuwen, P.C.J. Kamer, J.N.H. Reek, P. Dierkes, Chem. Rev. 100 (2000) 2741;
Z. Freixa, P.W.N.M. van Leeuwen, Dalton Trans. (2003) 1890.

[40] N.M. Scott, S.P. Nolan, Eur. J. Inorg. Chem. (2005) 1815.
[41] H. Jacobsen, A. Correa, A. Poater, C. Costabile, L. Cavallo, Coord. Chem. Rev., this issue.
[42] C. Chatfield, A.J. Collins, Introduction to Multivariate Analysis, Chapman and Hall Ltd., London, 1980;
J. Townend, Practical Statistics for Environmental and Biological Scientists, John Wiley & Sons Ltd., Chichester, 2002.
[43] D. Livingstone, Data Analysis for Chemists, Oxford University Press, Oxford, 1995.
[44] K. Hansch, A. Leo, R. Taft, Chem. Rev. 91 (1991) 165.
[45] E.M. Thorsteinson, F. Basolo, J. Am. Chem. Soc. 88 (1966) 3929.
[46] W.A. Henderson Jr., C.A. Streuli, J. Am. Chem. Soc. (1960) 5791.
[47] T. Allman, R.G. Goel, Canad. J. Chem. Rev. Canad. Chim. 60 (1982) 716.
[48] T.E. Müller, D.M.P. Mingos, Trans. Met. Chem. 20 (1995) 533;
J.M. Smith, N.J. Coville, L.M. Cook, J.C.A. Boeyens, Organometallics 19 (2000) 5273;
J.M. Smith, N.J. Coville, Organometallics 20 (2001) 1210.
[49] H.M. Senn, D.V. Deubel, P.E. Blöchl, A. Togni, G. Frenking, J. Mol. Struct. (Theochem) 506 (2000) 233.
[50] D. White, B.C. Taverner, P.G.L. Leach, N.J. Coville, J. Comput. Chem. 14 (1993) 1042;
D. White, C.B. Tavener, P.G.L. Leach, N.J. Coville, J. Organomet. Chem. 478 (1994) 205;
B.C. Taverner, J. Comput. Chem. 17 (1996) 1612;
A. Immirzi, A. Musco, Inorg. Chim. Acta 25 (1977) L41;
I.A. Guzei, M. Wendt, Dalton Trans. (2006) 3991.
[51] D. White, C.B. Taverner, N.J. Coville, P.W. Wade, J. Organomet. Chem. 495 (1995) 41.
[52] J.D. Smith, J.D. Oliver, Inorg. Chem. 17 (1978) 2585;
G. Ferguson, P.J. Roberts, E.C. Alyea, M. Khan, Inorg. Chem. 17 (1978) 2965.
[53] T. Bartik, T. Himmler, H.-G. Schulte, K. Seevogel, J. Organomet. Chem. 272 (1984) 29.
[54] T.L. Brown, Inorg. Chem. 31 (1992) 1286.
[55] G.M. Bodner, M.P. May, L.E. McKinney, Inorg. Chem. 19 (1980) 1951.
[56] B.E. Mann, C. Masters, B.L. Shaw, J. Chem. Soc. A (1971) 1104.
[57] J. Halpern, P.F. Phelan, J. Am. Chem. Soc. 94 (1972) 1881.
[58] G.M. Bodner, Inorg. Chem. 14 (1975) 1932.
[59] L.E. Manzer, C.A. Tolman, J. Am. Chem. Soc. 97 (1975) 1955.
[60] T.E. Barder, M.R. Biscoe, S.L. Buchwald, Organometallics 26 (2007) 2183.
[61] T.E. Barder, S.L. Buchwald, J. Am. Chem. Soc. 129 (2007) 12003.
[62] J.T. Desanto, J.A. Mosbo, B.N. Storhoff, P.L. Bock, R.E. Bloss, Inorg. Chem. 19 (1980) 3086.
[63] M.L. Caffery, T.L. Brown, Inorg. Chem. 30 (1991) 3907.
[64] K.J. Lee, T.L. Brown, Inorg. Chem. 31 (1992) 289;
M.G. Choi, T.L. Brown, Inorg. Chem. 32 (1993) 5603.
[65] M.G. Choi, T.L. Brown, Inorg. Chem. 32 (1993) 1548;
G. Choi, D. White, T.L. Brown, Inorg. Chem. 33 (1994) 5591;
D.P. White, T.L. Brown, Inorg. Chem. 34 (1995) 2718.
[66] R.J. Bubel, W. Douglass, D.P. White, J. Comput. Chem. 21 (2000) 239.
[67] B.J. Dunne, R.B. Morris, A.G. Orpen, J. Chem. Soc. Dalton Trans. (1991) 653.
[68] K.D. Cooney, T.R. Cundari, N.W. Hoffman, K.A. Pittard, M.D. Temple, Y. Zhao, J. Am. Chem. Soc. 125 (2003) 4318.
[69] C.H. Suresh, Inorg. Chem. 45 (2006) 4982.
[70] C.H. Suresh, N. Koga, Inorg. Chem. 41 (2002) 1573.
[71] T. Vreven, K. Morukuma, J. Comput. Chem. 21 (2000) 1419;
S. Dapprich, I. Komaromi, K.S. Byun, K. Morukuma, M.J. Frisch, J. Mol. Struct. (Theochem) 461/462 (1999) 1.
[72] J. Mathew, T. Tinto, C.H. Suresh, Inorg. Chem. 46 (2007) 10800.
[73] A.G. Orpen, Acta Crystallogr. B 58 (2002) 398, and reference cited;
F.H. Allen, Acta Crystallogr. B 58 (2002) 380.
[74] N. Fey, A. Tsipis, S.E. Harris, J.N. Harvey, A.G. Orpen, R.A. Mansson, Chem. Eur. J. 12 (2006) 291.
[75] R.A. Mansson, A.H. Welsh, N. Fey, A.G. Orpen, J. Chem. Inf. Model. 46 (2006) 2591, and references cited.
[76] N. Fey, J.N. Harvey, G.C. Lloyd-Jones, P. Murray, A.G. Orpen, R. Osborne, M. Purdie, Organometallics 27 (2008) 1372.
[77] A.B.P. Lever, Inorg. Chem. 29 (1990) 1271.
[78] A.B.P. Lever, Inorg. Chem. 30 (1991) 1980.
[79] A.M. Gillespie, K.A. Pittard, T.R. Cundari, D.P. White, Internet Electr. J. Mol. Des. 1 (2002) 242.
[80] G. Frenking, N. Fröhlich, Chem. Rev. 100 (2000) 717;
G. Frenking, K. Wichmann, N. Fröhlich, C. Loschen, M. Lein, J. Frunzke, V.M. Rayon, Coord. Chem. Rev. 238 (2003) 55.
[81] J. Chatt, A.A. Williams, J. Chem. Soc. (1951) 3061.
[82] J. Chatt, R.G. Wilkins, J. Chem. Soc. (1952) 273.
[83] D.S. Marynick, J. Am. Chem. Soc. 106 (1984) 4064.
[84] D.G. Gilheany, Chem. Rev. 94 (1994) 1339.
[85] T. Leyssens, D. Peeters, A.G. Orpen, J.N. Harvey, New J. Chem. 29 (2005) 1424.
[86] A.G. Orpen, N.G. Connelly, Organometallics 9 (1990) 1206.
[87] J. Kapp, C. Schade, A.M. El Nahasa, P.V. Schleyer, Angew. Chem. Int. Ed. Engl. 35 (1996) 2236.
[88] K. Kitaura, K. Morokuma, Int. J. Quantum Chem. 10 (1976) 325;
K. Kitaura, S. Sakaki, K. Morokuma, Inorg. Chem. 20 (1981) 2292.
[89] T. Ziegler, A. Rauk, Theor. Chim. Acta 46 (1977) 1.
[90] P.S. Bagus, K. Hermann, C.W. Bauschlicher, J. Chem. Phys. 80 (1984) 4378;
K. Hermann, P.S. Bagus, C.W. Bauschlicher, Phys. Rev. B 31 (1985) 6371.
[91] M. Mitoraj, A. Michalak, J. Mol. Model. 13 (2007) 347;
A. Michalak, M. Mitoraj, T. Ziegler, J. Phys. Chem. A 112 (2008) 1933.
[92] G. Frenking, K. Wichmann, N. Fröhlich, J. Grobe, W. Golla, D. Le Van, B. Krebs, M. Lage, Organometallics 21 (2002) 2921.
[93] T. Leyssens, D. Peeters, A.G. Orpen, J.N. Harvey, Organometallics 26 (2007) 2637.
[94] G. Paccioni, P.S. Bagus, Inorg. Chem. 31 (1992) 4391.
[95] M. Mitoraj, A. Michalak, Organometallics 26 (2007) 6576.
[96] A.E. Reed, L.A. Curtiss, F. Weinhold, Chem. Rev. 88 (1988) 899.
[97] F. Weinhold, C.R. Landis, Valency and Bonding: A Natural Bond Orbital Donor–Acceptor Perspective, Cambridge University Press, Cambridge, 2005.
[98] C.A. Tolman, J. Am. Chem. Soc. 92 (1970) 2953.
[99] D.M. Hawkins, J. Chem. Inf. Comput. Sci. 44 (2004) 1.
[100] D.E. Morris, F. Basolo, J. Am. Chem. Soc. 90 (1968) 2531.
[101] H.-R. Bjørsvik, U.M. Hansen, R. Carlson, Acta Chem. Scand. 51 (1997) 733.
[102] E. Burello, G. Rothenberg, Adv. Synth. Catal. 345 (2003) 1334;
G. Rothenberg, S.C. Cruz, G.P.F. van Strijdonck, H.C.J. Hoefsloot, Adv. Synth. Catal. 346 (2004) 467;
E. Burello, P. Marion, J.-C. Galland, A. Chamard, G. Rothenberg, Adv. Synth. Catal. 347 (2005) 803;
E. Burello, G. Rothenberg, Adv. Synth. Catal. 347 (2005) 1969;
J.A. Hageman, J.A. Westerhuis, H.-W. Fruehauf, G. Rothenberg, Adv. Synth. Catal. 348 (2006) 361.
[103] E. Burello, D. Farrusseng, G. Rothenberg, Adv. Synth. Catal. 346 (2004) 1844.
[104] E. Burello, G. Rothenberg, Int. J. Mol. Sci. 7 (2006) 375.
[105] I.P. Beletskaya, A.V. Cheprakov, Chem. Rev. 100 (2000) 3009.
[106] J. Gasteiger (Ed.), Handbook of Chemoinformatics—From Data to Knowledge, Wiley-VCH, Weinheim, 2003.
[107] M.L. Drummond, B.G. Sumpter, Inorg. Chem. 46 (2007) 8613.
[108] H. Kubinyi, in: J. Gasteiger (Ed.), Handbook of Chemoinformatics—From Data to Knowledge, vol. 4, Wiley-VCH, Weinheim, 2003, p. 1555.
[109] W.E. Steinmetz, Quant. Struct.-Act. Relat. 15 (1996) 1.
[110] K.B. Lipkowitz, M.C. Kozlowski, Synlett (2003) 1547.
[111] R. Bosque, J. Sales, J. Chem. Inf. Comput. Sci. 41 (2001) 225.
[112] T.M. Trnka, R.H. Grubbs, Acc. Chem. Res. 34 (2001) 18;
A.H. Hoveyda, A.R. Zhugralin, Nature 450 (2007) 243.
[113] G. Occhipinti, H.R. Bjørsvik, V.R. Jensen, J. Am. Chem. Soc. 128 (2006) 6952.
[114] M. Sparta, K.J. Borve, V.R. Jensen, J. Am. Chem. Soc. 129 (2007) 8487.
[115] A. Beste, O. Kramer, A. Gerhard, G. Frenking, Eur. J. Inorg. Chem. (1999) 2037;
G. Frenking, J. Organomet. Chem. 635 (2001) 9;
M. Cases, G. Frenking, M. Duran, M. Sola, Organometallics 21 (2002) 4182, Sp. Iss. SI;
R. Dorta, E.D. Stevens, N.M. Scott, C. Costabile, L. Cavallo, C.D. Hoff, S.P. Nolan, J. Am. Chem. Soc. 127 (2005) 2485;
F. Bessac, G. Frenking, Inorg. Chem. 45 (2006) 6956;
R.A. Kelly III, H. Clavier, S. Giudice, N.M. Scott, E.D. Stevens, J. Bordner, I. Samardjiev, C.D. Hoff, L. Cavallo, S.P. Nolan, Organometallics 27 (2008)202.