

Adding Value to Crystallographically-Derived Knowledge Bases[†]

Natalie Fey,* Stephanie E. Harris, Jeremy N. Harvey, and A. Guy Orpen*

School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, U.K.

Received October 28, 2005

A protocol for the partially automated computational investigation of crystal structure geometries of transition-metal complexes with unusual/outlier structural features has been developed for application in an e-science context. This protocol not only is envisaged as a part of knowledge base software packages such as *Mogul* but can also be used to further analyze the results of database searches. The issues arising from automating the initial input generation and DFT optimization of complexes have been examined and a procedure for extracting additional knowledge “value” from the computational results is described. Potential problems/weaknesses arising from the choice of computational approach and from errors in the crystal structure refinement are discussed. A range of likely outcomes of applying this protocol to database mining results is illustrated, with representative examples identified for tetracoordinate transition-metal complexes and ligand fragments (terminal chloride, monodentate phosphorus(III), and primary amine ligands) with unusual metal–ligand bond lengths.

INTRODUCTION

Structural databases such as the Cambridge Structural Database (CSD)^{4,5} and the RCSB Protein Database (PDB)⁶ have become important sources of information on both molecular and supramolecular geometries. With the number of available crystal structures deposited rising exponentially,⁴ the robust derivation of chemical knowledge from these data and hence the development of knowledge bases is becoming a vital area of research; the contingent applications of the CSD in a variety of disciplines have been extensively reviewed.^{7–9}

Software tools for database “mining”, i.e. the extraction and analysis of intra- and intermolecular structural data, continue to be developed.^{2,3,8,9} Expectations have changed with respect to the provision of electronic information and immediate results are often required.⁹ In recognition of these changes, two recent examples have moved beyond the immediate processing of database searches and are aimed at the compilation of knowledge bases, i.e. preprocessed libraries of structural/chemical information. *IsoStar*^{9,10} combines intermolecular nonbonded contact information extracted from both the CSD and PDB with calculated data, while *Mogul*^{2,9} is aimed at collating intramolecular geometry information in the CSD. Both knowledge bases allow *efficient* access to chemically well-defined information as they are based on a core set of preprocessed data. The chemical expertise embedded in these knowledge bases permits automatic updates in response to expansion of their parent database(s). In addition, they do not require the user to have expert knowledge of database mining and statistical analysis techniques, and the underlying data structure allows the

consideration of much larger and more diverse data sets than is feasible for manual searches.

The viability of any data survey, as opposed to studying individual molecular geometries, is dependent on the availability of structural data. Deposition of crystal structure data in publicly accessible databases relies principally on abstracting from peer-reviewed publications,⁹ a process which implies a selection process by authors in favor of reporting interesting and/or novel compounds. Despite the extensive use of single-crystal X-ray diffraction analysis to ascertain the outcome of chemical synthesis,⁵ and indeed despite the value of any crystal structure for expanding the scope of a data survey, much information thus remains inaccessible. The emerging field of e-science and the associated Grid architecture¹¹ seek to develop and exploit networked data storage and computing facilities. This is aimed at encouraging the electronic deposition of raw experimental data as well as managing database access and fostering the development of suitable statistical analysis techniques, the automated compilation of knowledge bases, and the utilization of distributed computing approaches.

The investigation of intramolecular structural parameters (molecular dimensions)^{2,5,9,12,13} is perhaps the most widespread use of structural databases. Researchers interested in e.g. transition-metal complexes often refer to surveys of crystal structure geometries to establish common structural features for specific metal ions or coordination geometries as well as periodic trends (see for example refs 5, 14, and 15). In addition, computational chemistry approaches have often been assessed by seeking to reproduce crystal structure trends or specific geometries,^{16–19} and parameters for empirical force fields have been derived from available structural data.²⁰ These database surveys are particularly useful, because the crystallographically observed molecular structure of many transition-metal complexes is directly and robustly related to their electronic structure, i.e. observed metal–ligand distances and coordination geometries can often be used to

* Corresponding author phone: +44 117 928 8158; fax: +44 117 925 1295; e-mail: Guy.Orpen@Bristol.ac.uk, (A.G.O.) and e-mail: Natalie.Fey@Bristol.ac.uk (N.F.).

[†] Development of a Ligand Knowledge Base, part 2. See ref 1 for part 1.

distinguish between different oxidation and spin states, and preferences can therefore be formulated for different metal centers. This robustness has been highlighted in the proposed development of a metallo-organic geometry library,³ which will be an extension of the current “organic” version of the *Mogul* software.²

While the primary aim of knowledge bases is the generalization of structural parameters, database mining often also allows the identification of outliers, i.e. structures that deviate from observed geometrical preferences and trends. Given the close link between molecular and electronic structure observed for transition-metal complexes (vide supra), genuine outliers can further our chemical insights, while the automatic detection and elimination of erroneous structures serves to protect the credibility of the parent structural database. Seeking explanations for outliers thus increases the knowledge value and robustness of both primary structural databases and their derived knowledge bases.

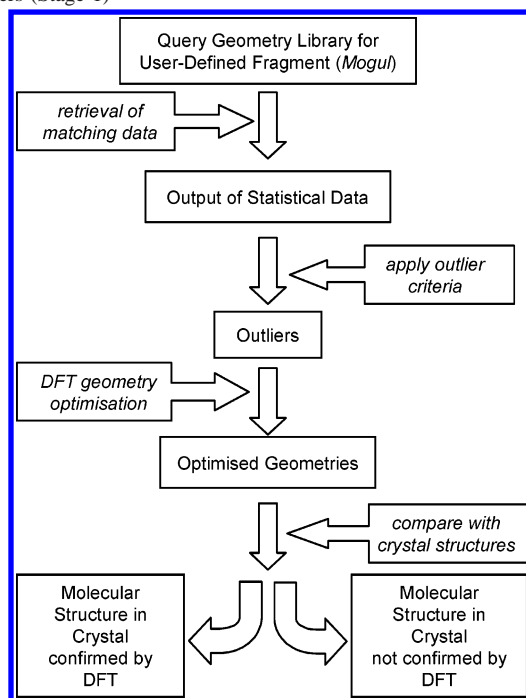
We have recently described the development of a prototype ligand knowledge base (LKB) for phosphorus(III) donor ligands.¹ A mature version of such a LKB will form a large-scale structural collection of knowledge about ligands found in metal complexes by combining information from crystallographic, computational, and possibly other sources. An understanding of both structural trends and outliers is vital in such an undertaking, and we have explored the scope for adding value to structural data “mined” from the CSD by calculations using density functional theory (DFT). Here we describe the sequence of steps allowing the partial automation of this process in an e-science context and illustrate possible outcomes with a series of representative case studies of coordination geometry and metal–ligand bonding. Challenges arising from automatic input generation and analysis as well as potential weaknesses of the chosen computational approach and the types of problems potentially arising from the original crystal structure determination are also discussed.

RESULTS AND DISCUSSION

(a) General Methodology. The widespread use of crystal structure databases such as the CSD and their derived knowledge bases crucially depends on their availability and ease of access.⁹ In line with this realization the process of adding value to a crystallographically derived ligand knowledge base by quantum mechanical (QM) or density functional theory (DFT) calculations should be amenable to at least partial automation. Assuming predefined criteria for outlier identification and structure comparison, Scheme 1 illustrates the sequence of steps that might be automated in the knowledge base development process (stage 1) derived from the *Mogul* software.^{2,3}

The process illustrated in Scheme 1 is initiated by querying a preprocessed knowledge base such as *Mogul*,^{2,3} which allows the rapid retrieval of data matching a query fragment. The current *Mogul* software uses a system of keys which describes the chemical environment of organic fragments in the CSD,² such that fragments with the same keys are considered chemically identical and stored together in a library. A similar geometry library can be developed for metal-based fragments (“*Metallo-organic Mogul*”),³ but different keys would be required to define the metal

Scheme 1. Automatic Identification and Computational Confirmation of Outliers (Stage 1)



environment (vide supra). After the standard *Mogul* procedure has been invoked to retrieve molecular geometry data for a query fragment, the built-in statistical analysis can be used to establish descriptive statistics for a structural parameter of interest as observed in the CSD, including its mean value, the sample standard deviation (SSD), and other descriptors of the distribution of values. It should be noted that the statistical data could also be compiled by manually formulated searches of a primary database such as the CSD.

If diagnostic criteria, e.g. for structures with parameter values more than 3–4 SSDs from their mean values, have been defined, these can be used to generate a list of outliers automatically. Alternatively, the user can manually identify structures for further investigation. A standard DFT geometry optimization protocol might then be applied and the resulting optimized geometries compared to the crystal structure-derived geometry. At the end of this automated procedure (stage 1), the outliers will be divided into two groups depending on whether the optimized geometry confirms the molecular geometry observed in the crystal structure or not. [Several factors need to be considered when comparing crystallographically observed and calculated geometries. In this work we are principally interested in the geometry around the transition-metal center, i.e. M–L bond lengths and L–M–L angles. More remote structural parameters, such as dihedral angles describing substituent conformations, have not been compared when assessing the reproduction of crystal structure geometries (cf. 1.iii). From database studies, the “inherent uncertainties” for crystal structures have been estimated as 0.01–0.02 Å and 1–2° for M–L distances and L–M–L angles, respectively.²¹ However, these effects are masked by the known overestimation of M–L distances by most modern density functionals (cf. 2.ii), and for our chosen computational approach (described in the Appendix) we have observed a RMS deviation of 0.048 Å and a mean absolute error of 0.042 Å for M–L distances in 27 optimized, mononuclear metallo-organic complexes (giving 47 M–L

distances, not including suspected cases of incorrect metal assignment) when compared to their crystallographically observed geometries. For most of the complexes discussed here, the reproduction of the coordination geometry within the angular ranges defined in the Database Mining section (Appendix) was the main criterion when assessing the calculated geometries. Deviations from the crystallographically observed M–L distances of up to 0.08 Å have been considered as acceptable. In cases where the calculated M–L distances were of particular interest, these results have been supplemented by more extensive database mining results and calculated energies.]

A number of challenges have to be considered in the implementation of this procedure, mainly arising from the automation of this verification approach in an e-science environment. These are summarized below.

(1.i) Identification of Outliers. To automate the analysis of statistical distributions, diagnostic criteria for outliers have to be defined. While simple criteria based on SSDs can be meaningful diagnostics, a priori classification of coordination geometries may be more difficult, because the distinction of borderline cases often requires detailed chemical understanding. The implementation of an automatic verification approach should therefore permit user intervention at this stage.

(1.ii) Generation of Input for Computational Confirmation. A default computational approach can be chosen and its performance assessed and verified ahead of implementation. To carry out the computation counterions, solvent molecules, etc. need to be removed and the spin multiplicity and molecular charge determined. These procedures in turn depend on the robust identification and categorization of ligands and their complexes. If the observed molecular geometry has been identified as an outlier because of an unusual electronic configuration, accurate charge and multiplicity assignments in an automated approach would require an extensive library of options to be compiled and coded. However, calculations based on the wrong electronic structure are likely to give readily observable discrepancies between the crystal structure molecular geometry and the DFT optimized result. These discrepancies will then be investigated at a later stage of the knowledge building process (vide infra).

(1.iii) Comparison of Calculated and Crystal Structure Molecular Geometries. A variety of options exist to assess whether the crystal structure molecular geometry has been confirmed by the calculated structure. Of these, root-mean-square (RMS) structural overlays of heavy atoms are popular for comparing molecules. However, they can fail to detect satisfactory agreement for a specific structural feature of interest, if e.g. conformational changes of remote substituents have occurred or metal–ligand bond lengths are systematically overestimated by the computational approach (as discussed below). Preset tolerances for a variety of structural parameters or a combination of approaches for comparison of molecular geometries may be more suitable and ideally a range of options should be made available.

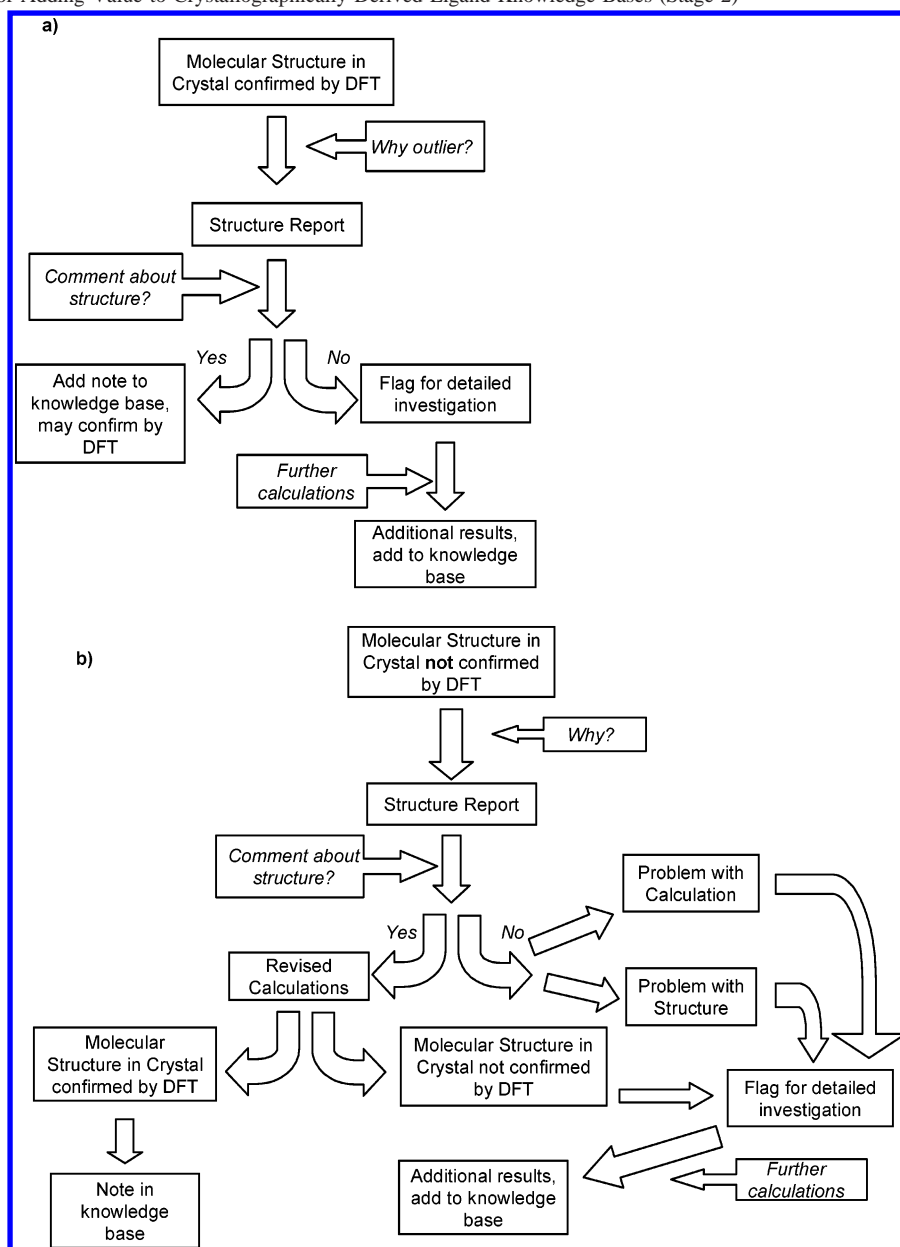
The integration of such a computational verification protocol in knowledge base generation software is particularly attractive because all outlier crystal structure molecular geometries are subjected to the same computational treatment. Considerable economies of scale thus offset the

challenges of automation, and this protocol could be implemented to allow a fully automated classification of observed outliers according to whether their geometries are or are not confirmed by computational chemistry.

However, capturing the chemical context of an unusual crystal structure molecular geometry by linking it with relevant publications and related experimental work can increase its “knowledge value” beyond simple computational verification. In addition, quantum mechanical (QM) and DFT calculation results contain further information (e.g. geometrical preferences, electronic structure, molecular properties), and, given the cost of performing calculations on large molecular systems at reasonably high levels of theory, exploitation of such additional “knowledge value” may be desirable. Scheme 2 formalizes the steps required for adding further value to knowledge bases (stage 2). Since different chemical contexts require considerable flexibility and chemical expertise to choose the most appropriate procedure for analysis and further investigation, the efficiency of a fully automated protocol will be limited, and we assume extensive user involvement for these value-added procedures. The scope for future automation, taking into account expected e-science developments, will be discussed in a later section (Summary and Outlook).

Frequently, unusual crystal structure molecular geometries are genuine outliers of chemical significance; notably in transition-metal complexes the ligands may have been designed to enforce unusual geometries. In this case the original structure report often includes a detailed discussion of the observed molecular structures, and, particularly in more recent publications, this may have been supplemented by additional experimental evidence, database surveys, and computational studies (vide infra). Crystal structure databases such as the CSD do not archive such information, so, as indicated in Scheme 2a,b, the literature structure report (and any associated discussions) should be consulted at an early stage of any value-added procedure; the subsequent referencing of existing discussions will enhance the value of both primary database and derived knowledge bases independent of further analysis of the calculated data. [Arguably this literature search step should be undertaken before any calculations are performed. However, it would be extremely difficult to integrate this step in the automated compilation of knowledge bases (stage 1, see Scheme 1), as it relies on an understanding of the literature and wider chemical context that is unlikely to be achieved by a computer program soon. Delaying this step until after the initial automatic procedure focuses the required user intervention on the procedure necessary for adding knowledge value.]

If the structure report or related publications provide a detailed explanation of the observed and calculated geometries (Scheme 2a, left), then it may suffice to augment the database entry and the derived knowledge base with a summary of this discussion and appropriate references. In some cases, modern computational chemistry results can provide more detailed confirmation and explanation of the structural feature of interest than was available at the time of original publication; this option will be illustrated in the case studies below (part b). On the other hand, if the structure report does not provide an explanation of the structural feature that led to its classification as an outlier, further analysis of the calculation results as well as more extensive

Scheme 2. Procedure for Adding Value to Crystallographically Derived Ligand Knowledge Bases (Stage 2)^a

^a Part (a), crystal structure molecular geometry confirmed by DFT, and part (b), crystal structure molecular geometry not confirmed by DFT.

calculations may be used to provide a rationalization (Scheme 2a, right).

The alternative scenario, where the calculated geometry does not confirm the observed crystal structure molecular geometry, has been summarized in Scheme 2b. Again the structure report may provide a straightforward explanation, such as a different electronic configuration from that assumed in the DFT optimization, and revised calculations will then allow the discrepancy to be resolved.

Alternatively, potential problems with either the calculated results or the crystal structure itself need to be considered to identify the likely reason for the observed divergence. The chosen computational approach needs to be suitable for the molecules of interest and its performance should be well established. Although we have used the popular B3LYP density functional with a fairly standard basis set combination in this study (see Computational Details, Appendix), other modern functional/basis set approaches would also be

suitable. [If few previous applications to transition-metal complexes have been reported for a chosen method, test calculations should be performed to confirm the suitability of the method used.] While other quantum mechanical approaches could also be used at this stage, we consider DFT the best present compromise between computational cost and structural and energetic accuracy for transition-metal complexes,²² given the computational resources currently available. However, standard computational approaches in general and specifically DFT have known weaknesses, which may cause discrepancies between calculated and crystallographically observed molecular geometries as listed below (2.i–2.iv). Point 2.i is generic across a range of computational methodologies, whereas the remaining points are specific to DFT. Some of the issues identified here will be further illustrated in part b.

(2.i) *Neglect of Complex Environment.* It is standard practice in computational chemistry to minimize computa-

tional cost by neglecting to take account of the crystal or solution environment and optimizing the complexes of interest in vacuo and in the absence of counterions. In many cases, the crystallographically observed conformation is at or near the calculated global energy minimum, and this simplification is justified.²³ If conformational energy differences are small, however, environmental effects may need to be taken into account. This may also be true in cases where the bonding within the 'molecular' complex being considered could also be described as involving ion pair character (see part b for an example).

(2.ii) *Systematic Overestimation of M–L Distances.* Most gradient corrected and hybrid density functionals are known to overestimate metal–ligand bond lengths systematically compared to geometries observed in crystal structures.^{16,18,19} This can be accommodated when comparing calculated and observed structures by setting suitable tolerances (vide supra, 1.i). If subtle structural changes occur in response to changes in the electronic structure (e.g. for different spin states) or metal–ligand distances for different metal centers are compared, further calculations may need to be performed to establish the size of this deviation for related complexes.

(2.iii) *Prediction of Spin State Energy Differences.* Comparison of calculated energies for different electronic configurations allows the determination of spin state energy preferences. If spin state effects are important, the performance of the chosen computational approach needs to be established, as many DFT functionals fail to predict energy differences between spin states accurately.^{18,19,24,25}

(2.iv) *Modeling of Weak Interactions.* Some molecular conformations observed in crystal structures are thought to be determined by weak, attractive inter- and intramolecular interactions (dispersion), e.g. H- π and π - π close contacts which are important in crystal packing. Most DFT functionals are not suitable for the study of dispersion.²⁶ However, the structural features of interest in the study of transition-metal complexes, i.e. metal–ligand bond lengths and coordination geometries, are unlikely to be determined by these weak interactions, and this problem will not be considered further here.

The discrepancy between observed and calculated molecular structure may also arise from problems with the reported crystal structure determination, and computational experiments can be used to investigate this further. In contrast to a computational model, where structural and energetic performance can be established by test calculations and problems/weaknesses are usually due to approximations made in the modeling approach, crystal structure refinement relies on the processing of experimental data. While the refinement process is well-established and generally robust, crystal structures with small and sometimes even large errors have been published.²⁷ Although many errors are identified and corrected either prior to publication or by the CCDC internal checking procedures,⁴ some structures with undetected problems are likely to be included in the CSD. In a recent paper Spek²⁷ discussed the importance of crystal structure validation (particularly when considering the large number of structures determined and published today) with reference to the program *PLATON*. The major sources of errors in crystal structure determinations are as follows:

(3.i) *Space Group Assignment.* In recent years there have been a number of publications reporting space group cor-

rections (see for example refs 28 and 29). Incorrect space group assignment may have a significant effect on bond lengths (and angles)²⁹ and so should be considered as a possible reason for unexplained outliers. Martin and Orpen²¹ used a version of *PLATON*³⁰ to check space group assignments for their database study and discarded ca. 1.5% of CSD entries with possible space group errors from their working data set. It is noted by Marsh that even with the development of checking routines, the number of assignments of space group Cc which are incorrect remains at 10% in the CSD.²⁹

(3.ii) *Cell Dimensions.* These may have been determined imprecisely if the crystal quality was low, but mistakes could also occur during data processing or in the compiling of structural data for publication.

(3.iii) *Undetected Disorder.* All the CSD searches discussed in this work have excluded disordered structures, but this relies on an explicit flag in the database. Outlier bond lengths and angles in particular may result from undetected or unresolved disorder. Checking of displacement ellipsoid plots or displacement parameters for large values or unusually large estimated standard deviations (ESDs) for atomic coordinates, bond lengths, and angles is routinely carried out by most crystallographers, as unusual displacement parameters may be indicative of other problems such as systematic errors in intensity data. However, if librational, thermal, or positional disorder is not discussed in the structure report, then its occurrence will have remained unnoticed. We have not come across any definite examples of undetected disorder in the database searches discussed in this work (see however ref 31), but consideration of this problem should be a routine part of the analysis of deviations between computed and crystallographically determined structures.

(3.iv) *Elemental Composition.* Errors of this sort include misassignment of the number of hydrogen atoms in a structure as well as of element types, for example, nitrogen versus oxygen²⁷ and, more importantly, the incorrect identification of metal atoms. Differences in the number of electrons may not be detected in the refinement of X-ray diffraction data, so an apparently acceptable crystal structure determination can be realized using an incorrect assignment of atom identities. However, the geometry of the metal coordination sphere is often robust to this error, and the DFT optimized complex may not reproduce the crystal structure molecular geometry and thus be flagged as a computationally unconfirmed outlier in the validation process described here. A combination of database mining and computational chemistry can be used to suggest alternatives; however, only a refinement of the original diffraction data with a corrected composition can offer more conclusive evidence.

In summary, a general methodology for the automated identification and computational verification of outliers in a knowledge base of metal–ligand fragments has been proposed (stage 1). In addition, protocols for (manually) adding value to the resulting knowledge base by investigation of the chemical/literature context and additional calculations have been described (stage 2) and the most likely sources of discrepancies between observed and calculated molecular structures have been discussed. The likely outcomes of this methodology will be illustrated by the case studies in part b below.

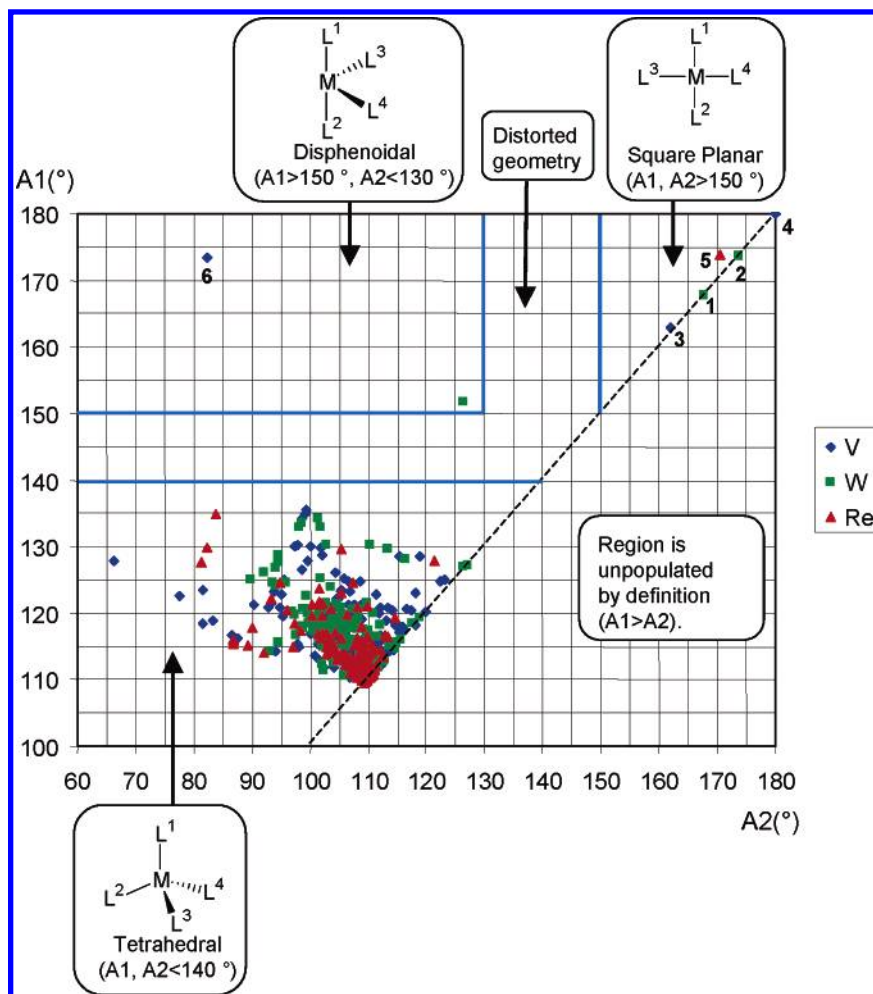


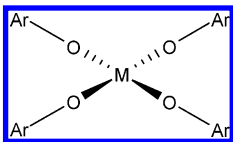
Figure 1. Distribution of tetracoordinate V, W, and Re fragments as defined by the maximum L^1-M-L^2 angle, $A1$ and L^3-M-L^4 angle, $A2$ (see Tables 1 and 2 for complex numbering).

(b) Case Studies. (i) *Molecular Structure in Crystal Confirmed by DFT (Scheme 2a).* With the continuing expansion of the CSD, recent crystal structure reports in particular are likely to include a discussion of unusual molecular structures in the context of database mining and computational results, and some representative examples are listed in the Supporting Information (Table S1). While this information is not presently indicated for CSD entries, it adds considerable knowledge value to both primary databases and derived knowledge bases, and the procedures outlined in part a are designed to facilitate the contextualization of crystallographically observed geometries. In addition, some crystal structure reports include a speculative discussion of an unusual structural parameter or bond length, which could be tested by reasonably straightforward (DFT) calculations. These have not been attempted in the original study, often an older publication where computational resources were probably limited at the time. Typical are studies of the aryloxide complexes $(M(OAr)_nL_{4-n})$ of vanadium, tungsten, and rhenium, which deviate from the approximately tetrahedral coordination geometry more usually observed for tetracoordinate complexes of these metals (Figure 1) due to an electronic preference for the square planar geometry. It should be noted that our analysis has focused on the broad classification of tetracoordinate crystal structure geometries as described in the Appendix and that outliers within the tetrahedral subset have not been investigated. Information about the metal

oxidation state and d-electronic configuration is not easily/automatically accessible from crystallographic data such as the CSD and has therefore not been included in Figure 1. Table 1 summarizes the relevant results (see Appendix for data mining and computational details). A more detailed description of the main structural features for these and all other DFT-optimized complexes discussed in this work may be found in the Supporting Information (Table S2).

As shown in Table 1b, the electronic preference for square planar coordination geometries observed for low spin, d^2 W(IV) and d^3 Re(IV) in **1**, **2**, and **5** arises from a relative stabilization of the d_{z^2} orbital in response to strong π -bonding across the MO_4 core.^{32,35,36} For the high spin, d^3 V(II) complexes **3** and **4**, a square planar coordination geometry gives rise to a higher ligand field stabilization energy than a tetrahedral geometry, and this may be further reinforced by partial V–O π -bonding interactions.^{33,34} DFT calculations allow the direct comparison of the energies for square planar and tetrahedral aryloxide complexes (Table 1c). We have optimized these coordination geometries for the W(DMP)₄ complex **2**, imposing D_4 and S_4 symmetry, respectively. With these symmetry constraints, the square planar (D_4) geometry is clearly energetically favored over the tetrahedral (S_4) (Table 1c). The formation of a delocalized π -system in the square planar geometry is further confirmed by the observed linearity of the C–O–W bonds. We have also fully optimized the geometry of complex **2** and the corresponding

Table 1. Summary of Results for Tetracoordinate Complexes of Vanadium, Tungsten, and Rhenium^a

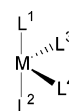
(a) Database Mining						
CSD search fragment				no. of fragments	summary of results	
4-coordinate M (M = V, W, Re), defined by A1 (max. L ¹ –M–L ²) and A2 (L ³ –M–L ⁴)				V: 370 W: 208 Re: 211	main geometry tetrahedral, a few square planar and disphenoidal exceptions ^b (Figure 1)	
(b) Description of Outliers						
						
no.	CSD Refcode	M	O–Ar	A1, A2/°	summary of structure report	ref
1	DETWII10	W(IV), d ²	DIPP ^c	168, 168	Extensive π -donation from oxygen LP to metal, delocalized π system over WO ₄ core, qualitative MO diagram shown. Sharp ¹ H NMR data for both complexes suggest diamagnetism.	32
2	FUWZUS	W(IV), d ²	DMP ^d	174, 174		
3	KELGOX	V(II), d ³	DIPP ^{c,e}	163, 162	Electronic preference of tetracoordinate d ³ complex for square planar geometry. Magnetic moment ($\mu_{\text{eff}} = 3.83 \mu_{\text{B}}$) consistent with $S = 3/2$.	33
4	PERZOB	V(II), d ³	O-2,6- ^t Bu ₂ -4-Me-C ₆ H ₂ ^f	180, 180	Partial V–O π -bond character explains enlarged V–O–C angle (168°). Magnetic moment ($\mu_{\text{eff}} = 3.71 \mu_{\text{B}}$) consistent with $S = 3/2$.	34
5	SAVKUV10	Re(IV), d ³	DIPP ^c	174, 171	Discussion of π interactions between M d- and O p-orbitals, preference for square planar due to π -delocalization and relative stabilization of Re d ₂₂ orbital. Paramagnetic ($\mu_{\text{eff}} = 1.50 \mu_{\text{B}}$).	35
(c) Calculation Results						
no.	input description			A1, A2/°	additional comments	
2	from XRD			178, 175	$S = 0$, see Figure 2 for delocalized π -orbital.	
	from tetrahedral			178, 178	returns to square planar	
	D_4 , square planar ^g			180, 180	rel $E = 0.0 \text{ kcal mol}^{-1}$, C–O–W = 180° ^h	
	S_4 , tetrahedral			115, 115	rel $E = 32.4 \text{ kcal mol}^{-1}$, C–O–W = 138° ^h	
3a	from XRD, DIPP simplified to DMP			164, 150	$S = 3/2$ ⁱ	
4	from XRD			180, 180	$S = 3/2$ ⁱ	
5a	from XRD, DIPP simplified to DMP			175, 174	$S = 1/2$	
	tetrahedral geometry			179, 179	returns to square planar	

^a See Appendix for further details. ^b Tetrahedral: A1, A2 < 140°; square planar: A1, A2 > 150°; disphenoidal: A1 > 150°; A2 < 130°. ^c DIPP = O-2,6-^tPr₂-C₆H₃. ^d DMP = O-2,6-Me₂-C₆H₃. ^e [V(DIPP)₄{Li(thf)}₂]. ^f Only 2 OAr ligands coordinated to V: [V(O-2,6-^tBu₂-4-Me-C₆H₂)₂py₂]. ^g Most similar to crystal structure molecular geometry. Complexes with D_{2h} and D_{4h} symmetry are higher in energy (by 6.9 and 3.4 kcal mol^{−1}, respectively) due to unfavorable interactions of methyl groups. ^h Preference for C–O–W angle shown was confirmed by single point energies. ⁱ High spin energetically favored for first row transition metal.

^a See Appendix for further details. ^b Tetrahedral: A1, A2 < 140°; square planar: A1, A2 > 150°; disphenoidal: A1 > 150°; A2 < 130°. ^c DIPP = O-2,6-^tPr₂-C₆H₃. ^d DMP = O-2,6-Me₂-C₆H₃. ^e [V(DIPP)₄{Li(thf)}₂]. ^f Only 2 OAr ligands coordinated to V: [V(O-2,6-^tBu₂-4-Me-C₆H₂)₂py₂]. ^g Most similar to crystal structure molecular geometry. Complexes with D_{2h} and D_{4h} symmetry are higher in energy (by 6.9 and 3.4 kcal mol^{−1}, respectively) due to unfavorable interactions of methyl groups. ^h Preference for C–O–W angle shown was confirmed by single point energies. ⁱ High spin energetically favored for first row transition metal.

Re(DMP)₄ complex **5a** from a tetrahedral starting geometry derived from that observed for the related d⁰ [Ti(IV)DIPP₄] complex (CSD Refcode DOFWOK).³⁷ Both **2** and **5a** return to square planar geometries. In addition, the preference for a square planar coordination has been confirmed for the vanadium complexes **3a** and **4** (Table 1c).

In addition to comparing different coordination geometries energetically, the impact of steric and/or agostic interactions on a crystallographically observed coordination geometry can be investigated by performing computational “experiments”, where the substituent of interest is replaced by an alternative (usually smaller) group, and structural changes are monitored. [If the observed complex geometry is indeed a result of steric/agostic interactions, replacing the relevant bulky group with a smaller substituent should give rise to a change in geometry, usually involving relaxation toward the geometry more commonly observed for complexes of the same metal, oxidation, and spin state. However, if electronic effects contribute to the stabilization of the observed geometry, then no significant structural change would be expected on removal of the interacting group(s). It should be noted that

Chart 1. Disphenoidal/Sawhorse Geometry

this simple test can fail if changing the substituent significantly alters the electronic properties of the ligand or conformational changes give rise to other interactions, so further studies may be necessary.] A number of tetracoordinate complexes in the CSD adopt disphenoidal geometries (Chart 1, A1 (L¹–M–L²) > 150°, A2 (L³–M–L⁴) < 130°) that can also be described as sawhorse or *cis*-divacant octahedral, with the vacant sites occupied by intramolecular agostic interactions and/or blocked by bulky ligand substituents. Table 2b lists some examples, where this interpretation was suggested in the original structure report. Alvarez¹⁴ presented a more detailed discussion of this “sawhorse” coordination geometry and its chemical significance; the order of d-orbital energy levels in such species has also been discussed.³⁸

Table 2. Summary of Results for Tetracoordinate Complexes of Vanadium and Ruthenium Adopting Disphenoidal Geometries

(a) Summary of Database Search						
CSD search fragment		no. of fragments	summary of results ^a			
4-coord. V		370	1 disphenoidal geometry (6)			
4-coord. Ru		22	[2 square planar (3 , 4), 367 tetrahedral cf. Figure 1, Table 1a] 5 disphenoidal geometries ^b [3 square planar, 13 tetrahedral, 1 distorted]			
(b) Description of Outliers						
no.	CSD Refcode	complex	A1, A2/°	M–C _{agostic} /Å	summary of structure report	ref
6	PERZUH	V(O-2,6-Ph ₂ –C ₆ H ₃) ₂ (tmeda) ^c	173, 82	2.66, 2.72	Short V···H–C _{Ph} distance for Ph rings adjacent to vacant site, due to bulky groups. Magnetic moment ($\mu_{\text{eff}} = 3.63\mu_{\text{B}}$) consistent with $S = 3/2$.	34
7	GOYGOQ ^d	RuCl ₂ (PPh ₂ (2,6-Me ₂ –C ₆ H ₃) ₂	168, 102	2.65, 2.65	2 agostic Ru···H–C interactions, steric protection of 14-electron complex by bulky PR ₃ . Diamagnetic.	39
8	CEPQIX	[RuPh(CO)(P ^t Bu ₂ Me) ₂] BAR' ₄ ^e	168, 94	2.88, 2.87	Ph, CO cis, PR ₃ trans. Vacant sites occupied by agostic C–H bonds from ^t Bu on different PR ₃ . Addt. IR evidence. NMR data suggests diamagnetism.	40
9	VENBUL ^d	[Ru(CH=C(SiMe ₃)(Ph))(CO)(P ^t Bu ₂ Me) ₂]BAR' ₄ ^e	171/165, 93/88 ^f	3.05/3.09, 3.74/3.47 ^f	1 short agostic donation from Ph, 1 longer from ^t Bu group. Additional spectroscopic evidence. NMR data suggests diamagnetism.	41
10	FOSHUQ10	[Ru(SC ₆ F ₅) ₂ (PPh ₃) ₂]	158, 106	3.28, 3.29	SAr trans, PR ₃ cis. Distorted octahedron with vacant sites occupied by H from ortho-C of Ph group. No indication of spin state/magnetic moment given.	42
(c) Calculation Results						
no.	input description		A1, A2/°	M–C _{agostic} /Å	comments	
6	from XRD		170, 82	2.74, 2.78	$S = 3/2^g$	
6a	V(O-2-Ph-6-Me-C ₆ H ₃) ₂ (tmeda), simplified 6		118, 80	(3.58, 3.78)	Replaced Ph rings involved in agostic interaction by Me groups. N–V–N angle virtually unchanged as rigid tmeda. $S = 3/2^g$	
7	from XRD		167, 105	2.66, 2.64	Ru–P–C _{xy1} –C _{Me,ag} = 4 and 6°	
7a	RuCl ₂ (PPh ₂ (2-Me-C ₆ H ₄) ₂ , simplified 7		143, 99	(3.46, 3.48)	Replaced Ph rings involved in agostic interaction by Me groups. Ru–P–C _{Ph} –C _{Ph} = 40 and 39°.	
8	from XRD		164, 93	3.11, 3.16		
8a	[RuPh(CO)(PMe ₃) ₂] ⁺ , simplified 8		174, 97	(3.56, 3.58)	rel $E = 0.0 \text{ kcal mol}^{-1}$	
	tetrahedral		175, 97		returns to disphenoidal	
	square planar		178, 167		rel $E = 33.8 \text{ kcal mol}^{-1}$. Ph ring deformed from planarity.	
^a Geometries defined by A1 and A2 as shown in Figure 1 and Table 1a. ^b Ru(II) complexes summarized in Table 2b, see also Table S1 for distorted Ru(0) complex and Tables 4b and 8b and S1 for square planar complexes. ^c tmeda = <i>N,N,N,N</i> -tetramethylethylenediamine. ^d In the recent (November 2004) version of the CSD structures 7 and 9 are represented with the C–H agostic interactions described, making the ruthenium formally 6- and 5-coordinate, respectively. ^e Ar'=(3,5-(F ₃ C) ₂ –C ₆ H ₃). ^f Two independent molecules in unit cell. ^g High spin energetically favored for first row transition metal.						

^a Geometries defined by A1 and A2 as shown in Figure 1 and Table 1a. ^b Ru(II) complexes summarized in Table 2b, see also Table S1 for distorted Ru(0) complex and Tables 4b and 8b and S1 for square planar complexes. ^c tmeda = *N,N,N,N*-tetramethylethylenediamine. ^d In the recent (November 2004) version of the CSD structures **7** and **9** are represented with the C–H agostic interactions described, making the ruthenium formally 6- and 5-coordinate, respectively. ^e Ar' = (3,5-(F₃C)₂–C₆H₃). ^f Two independent molecules in unit cell. ^g High spin energetically favored for first row transition metal.

To illustrate the additional value generated from such computational “experiments”, we have investigated complexes **6–8** in detail. In the absence of π -delocalized electronic stabilization (vide supra), a tetracoordinate vanadium complex is likely to adopt a tetrahedral geometry (Figure 1, Tables 1 and 2).^{14,43} This prediction is confirmed by the computed geometry for complex **6a** (Table 2c), where the phenyl substituents pointing toward the vacant sites on the vanadium center have been replaced by hydrogen atoms. Removal of steric bulk from **6** thus results in considerable structural relaxation toward a tetrahedral geometry, even though the rigid tmeda ligand prevents a relaxation of the N–V–N angle.

For the ruthenium(II) complexes **7–10** geometry predictions are less straightforward. With only 22 tetracoordinate ruthenium fragments recorded in the CSD (Table 2a), no clear structural preference can be determined. Based on shape measures Alvarez and co-workers¹⁴ found nearly tetrahedral

and square planar geometries to be more common than the disphenoidal/sawhorse geometry for tetracoordinate d⁶ metal complexes in general. An experimental investigation of tetracoordinate d⁶ iron(II) complexes of the general form L₂FeX₂ indicated the influence of different spin states as well as ligand steric and electronic effects on the geometrical preferences;⁴⁴ only tetrahedral and square planar geometries are observed. Poli⁴³ suggested a triplet d⁶ square planar configuration as most likely for open-shell 4d and 5d metals. Complexes **7** and **8** have the shortest Ru...C_{agostic} distances of the ruthenium complexes shown in Table 2, and the structure reports present additional experimental evidence and detailed discussions of agostic and steric interactions. Both complexes appear to be diamagnetic, in light of the ¹H NMR data reported.

After the removal of the phenyl substituents involved in agostic interactions, optimized **7a** shows some relaxation toward a tetrahedral geometry (Table 2c). In addition, the

Table 3. Summary of Results for Fe(III) Complexes with Monodentate Tertiary Phosphine (PA₃) Ligands (**11**, **12**)

(a) Summary of Database Search						
CSD search fragment			no. of fragments		summary of results	
Fe(III)–PA ₃ , A = any type of C, P–C bonds acyclic			8		Fe–P range: 2.229–2.655 Å, single outlier (SAVHIG, 11) with long Fe–P (2.655, 2.622 Å)	
P–Fe–PA ₃ (P ligands <i>trans</i> : P–Fe–P = 150–180°)			125		Fe–PA ₃ range: 2.155–2.361 Å except outlier, 11	
5-coordinate Fe(III)-Cl			62		mean Fe(III)–Cl = 2.247(43) Å ³	
(b) FeX ₃ (PA ₃) ₂ Complexes in CSD						
no.	CSD Refcode	complex	Fe–X/Å	Fe–P/Å	summary of structure report	ref
11	SAVHIG	[FeCl ₃ (PPh ₃) ₂]	2.216, 2.207, 2.195	2.655, 2.622	high spin (<i>S</i> = 5/2), EPR, UV/vis and magnetic susceptibility data	51
12	SAVHOM	[FeCl ₃ (PMe ₃) ₂]	2.231, 2.231, 2.232	2.341, 2.332	thermal equilibrium between <i>S</i> = 5/2 and <i>S</i> = 3/2, <i>S</i> = 3/2 ground state. EPR silent, but UV/vis and magnetic susceptibility data	52
13	JODFOX	[FeBr ₃ (PMe ₂ Ph) ₂]	2.350, 2.394, 2.413	2.343, 2.345	not discussed	
(c) Calculation Results						
no.	<i>S</i>	Fe–Cl/Å	Fe–P/Å	rel <i>E</i> /kcal mol ^{–1}	comment	
11	1/2	2.209, 2.241, 2.220	2.385, 2.379	28.2	<i>S</i> = 5/2	
	3/2	2.304, 2.265, 2.299	2.387, 2.393	8.3		
	5/2	2.260, 2.262, 2.271	2.666, 2.670	0.0		
	expt	2.216, 2.207, 2.195	2.655, 2.622			
12	1/2	2.225, 2.255, 2.225	2.339, 2.334	21.8	<i>S</i> = 3/2 & 5/2 equilibrium, <i>S</i> = 3/2 ground state	
	3/2	2.306, 2.276, 2.306	2.349, 2.348	1.0		
	5/2	2.278, 2.278, 2.296	2.586, 2.587	0.0		
	expt	2.231, 2.231, 2.232	2.341, 2.332			

orientation of the phenyl rings changes from being almost coplanar with the M–P bonds to a more staggered arrangement. However, the overall structural change is much less pronounced than that observed for the vanadium complex **6a**, suggesting that electronic effects may contribute to stabilize the observed geometry (discussed below). The agostic interactions in complex **7** have recently been investigated by neutron diffraction, NMR spectroscopy, and DFT calculations.⁴⁵ The calculations were used to elucidate the electronic structure of a simplified model complex, illustrating the electronic stabilization of the disphenoidal geometry. In complex **8a**, where the *tert*-butyl groups have been replaced by methyl substituents, the DFT optimized geometry maintains the disphenoidal geometry observed for **8** (Table 2c). Optimization of **8a** from a tetrahedral starting geometry returns to the disphenoidal complex, while a square planar complex with the phosphine ligands remaining in a *trans* configuration is significantly higher in energy (Table 2c).

Other computational studies have confirmed the disphenoidal geometry as favorable for electron deficient ruthenium(II) and iridium(III) complexes,^{46,47} even in the absence of bulky ligands and agostic interactions. Disphenoidal geometries have also been predicted for d⁶ M(CO)₄ species.^{38,48} This geometry is thought to be favored over square planar coordination, because it is associated with a larger HOMO–LUMO gap⁴⁶ due to a second-order Jahn–Teller effect.⁴⁹ For complex **8a**, the disphenoidal geometry observed crystallographically for **8** is thus not occurring solely in response to steric and/or agostic interactions but is electronically favorable. However, the latter interactions are also important in complexes **7** and **8**, because they prevent dimerization or solvent coordination for these 14-electron ruthenium(II) complexes, thereby allowing their isolation and characterization.⁵⁰

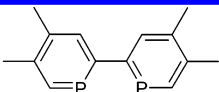
(ii) *Molecular Structure in Crystal Not Confirmed by DFT* (Scheme 2b). If the crystal structure molecular geometry is not verified by the DFT optimization outcome, further investigation is necessary to discover the cause of this discrepancy. The most straightforward explanation relates to the electronic structure of the complex (i.e. spin state and/or oxidation state, Scheme 2b and part a, 1.ii) and in many cases, a calculation with the correct electronic configuration will be able to resolve the discrepancy. However, the known failure of most DFT approaches to predict small spin state energy differences accurately (part a, 2.iii) may need to be taken into account when assessing structural agreement and energetic preferences. As an illustration of this point, Table 3 summarizes the results for Fe(III) complexes with monodentate tertiary phosphines.

There are few Fe(III) complexes with monodentate tertiary phosphines (PA₃) in the CSD (Table 3a) due to the low affinity of hard metal centers such as Fe(III) for soft P-donor ligands. As shown in Table 3a,b the trigonal bipyramidal complex, [FeCl₃(PPh₃)₂] **11**, has two long Fe–P bond lengths (2.655, 2.622 Å) which are clear outliers compared to other Fe–PR₃ complexes, including the analogous [FeCl₃(PMe₃)₂] **12** (Fe–P 2.341, 2.332 Å). A variation in the phosphine ligand substituents as seen for complexes **11** and **12** (cf. phenyl versus methyl groups) would generally be considered as an “innocent” ligand derivation and so would be expected to have minimal structural and electronic consequences. The combination of the strong field phosphine ligands and the weak field chlorides makes a prediction of spin state preferences difficult without further experimental data, but an automated approach would likely assume the same spin state for both complexes.

In practice the observed difference in Fe–P bond lengths for **11** and **12** suggests different spin states of the d⁵ Fe(III) metal center; indeed this was confirmed in the structure report

Table 4. Summary of Results for Ruthenium(–II) Biphosphinine Complexes **14**^a

(a) Description of Complex



tetramethylbiphosphinine (tmbp)

no.	CSD Refcode	complex	A1, A2/°	summary of structure report	ref
14	QOZMEX	[Ru(tmbp) ₂].[Li(thf) ₃] ₂	175, 175	<p>Long Ru...Li contacts (2.740(3) Å) interpreted as weak, mainly electrostatic interaction. If tmbp neutral and electronically innocent, Ru(–II), d¹⁰, expected to be tetrahedral. Square planar geometry explained by tmbp^{•–} radical anions coordinating d⁸ metal center. Extended Hückel calculations to illustrate electronic structure.</p> <p>Later publication highlighted importance of counterion for favoring square planar geometry, small energy difference between tetrahedral and square planar geometry for d¹⁰ [M(tmbp)₂]^q complexes from DFT conformational study. Confirmed different formal electron counts in square planar and tetrahedral geometry.</p>	54 55

(b) Calculation Results

no.	input description	A1, A2/°	tmbp...tmbp dihedral/°	comments
14	from XRD, no counterion	154, 154	40	rel <i>E</i> = 0.1 kcal mol ^{–1}
	tetrahedral	128, 127	80	rel <i>E</i> = 0.0 kcal mol ^{–1}
14a	[Ru(bp) ₂] ^{2–} , removed Me groups	148, 148	49	rel <i>E</i> = 0.04 kcal mol ^{–1}
	from tmbp ligands, XRD			
	tetrahedral	129, 127	79	rel <i>E</i> = 0.0 kcal mol ^{–1}
	square planar, <i>D</i> _{2h}	180, 180	0	rel <i>E</i> = 2.4 kcal mol ^{–1}
	broken symmetry	172, 172	12	rel <i>E</i> = 42.6 kcal mol ^{–1} , tmbp not planar
	broken symmetry, <i>D</i> _{2h}	180, 180	0	rel <i>E</i> = 50.5 kcal mol ^{–1}
14b	[Ru(tmbp) ₂].[Li(thf) ₃] ₂	167, 166	21	Ru...Li = 2.751, 2.754 Å

^a See Table 2a for results of database search for tetracoordinate ruthenium.

^a See Table 2a for results of database search for tetracoordinate ruthenium.

by magnetic susceptibility measurements, EPR and UV/vis spectroscopy, and qualitative molecular orbital considerations (Table 3b). The triphenylphosphine complex **11** adopts a high spin (*S* = 5/2) electronic configuration, whereas the experimental evidence for the trimethylphosphine complex **12** is consistent with an equilibrium between intermediate (*S* = 3/2) and high spin (*S* = 5/2) configurations, with a quartet ground state. Table 3c summarizes the calculated M–L bond lengths for low, intermediate, and high spin electronic configurations of complexes **11** and **12** along with their relative energies and experimentally observed spin state preferences. The preference of **11** for a high spin (*S* = 5/2) configuration is clearly confirmed, and the difference in bond lengths observed in the crystal structure geometries is apparently a result of the different electronic configurations of **11** and **12**. However, for the trimethylphosphine complex **12** the results also highlight that, when considered in isolation, neither the calculated bond lengths nor the calculated energy differences are sufficient to confirm the experimentally observed spin states unambiguously. Specifically, the Fe–P distance for the low and intermediate spin states are quite similar, but the intermediate spin is energetically favored (Table 3c). On the other hand, the energy difference between intermediate and high spin is very small, but the Fe–P distances for the intermediate spin configuration are in much better agreement with the crystal structure molecular geometry (Table 3b), suggesting that our chosen DFT approach slightly overestimates the energetic preference for the highest spin state (vide supra).

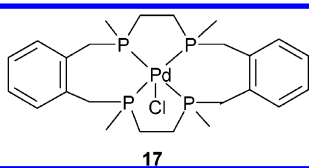
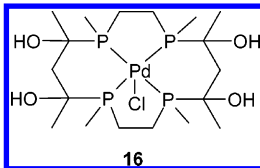
Complexes of iron(II) and iron(III), particularly those with nitrogen donor ligands, are among the most likely examples for differing spin state energy preferences^{19,25} and as such

would merit special attention and warning flags in an automated approach. We have previously reported³ how high and low spin states for Fe(II/III) 6-coordinate pyridine complexes may be identified by the bimodal distributions of Fe–N bond lengths in the CSD, confirming that both spin states can be observed experimentally. However, energetic preferences for most transition-metal complexes are much more distinct, and a qualitatively correct energetic ordering can usually be calculated.

The electronic structure of some ligands can also vary in response to their coordination environment, and complexes of nitrosyl (NO) are probably the most prolific examples of electronically noninnocent ligands in the CSD. Their coordination chemistry has been discussed extensively (see for example ref 53), and they will thus not be examined further except for noting that the M–N–O angle often allows distinction between NO⁺ and NO[–], with observed bond angles ca. 160° and 125°, respectively. An automated approach could flag complexes with NO ligands as potentially problematic and require additional user intervention; for most other electronically noninnocent ligands their electronic structure is likely to be of interest to the authors and thus discussed in the original structure report.

While it is common practice in computational chemistry to neglect the crystal environment (part a, 2.i), structural verification by DFT may fail if the environment influences the crystallographically observed molecular structure. This is most likely to occur if conformational energy differences are small or adducts with long M–L bonds are formed, i.e. where intramolecular interactions are weak; analysis of two such examples is shown in Tables 4 and 5.

Table 5. Summary of Results Relevant to $\text{PdCl}_2(\text{PMe}_2\text{Ph})_3$ Complex **15**

(a) Summary of Database Search				
CSD search fragment	no. of fragments	summary of results		
Pd(II)–Cl	1215	mean Pd–Cl ³ = 2.330(52) Å		
5-coord Pd(II)–Cl	13, cf. 1200 4-coord ³	range: 2.25–2.50 Å (3 outliers excluded) 3 outliers with long (>2.8 Å) Pd–Cl		
(b) Description of Outliers				
<div></div>				
no.	CSD refcode/complex	Pd–Cl/Å	summary of structure report/comment	ref
15	MPCPDS/PdCl ₂ (PMe ₂ Ph) ₃	2.956, 2.434	Long Pd–Cl bond noted, discussion of tight ion pair vs 5-coordination. UV/vis results favoring 5-coordination, intermediate between trigonal bipyramidal and square pyramidal ($\tau=0.28$; ^a $\alpha=\text{max}$. P–Pd–Cl=150°, $\beta=\text{max}$. P–Pd–P=167°).	57
16	CEFYIV	2.831	tetradentate P ligands restrict coordination geometry to square pyramidal with long axial bond	58
17	DOKWEF	3.106		59
(c) Calculation Results				
no.	input description	Pd–Cl/Å	comments	
15	from XRD	2.703, 2.559	Closer to trigonal bipyramidal ($\tau=0.40$; $\alpha=\text{max}$. P–Pd–Cl=141°, $\beta=\text{max}$. P–Pd–P=165°), both Cl in equatorial plane. NBO charges: Cl _{ax} –0.734, Cl _{eq} –0.659.	
15a	[PdCl(PMe ₂ Ph) ₃] ⁺ (removed axial Cl)	2.378	square planar (A1=178°, A2=170°)	
15b	[PdCl ₂ (PMe ₂ Ph) ₃]·CH ₂ Cl ₂ (close contact in crystal structure)	2.796, 2.508	($\tau=0.37$; $\alpha=\text{max}$. P–Pd–Cl=143°, $\beta=\text{max}$. P–Pd–P=165°)	
15 _{aq}	from XRD, in water continuum dielectric field	2.970, 2.511	closer to square pyramidal geometry than 15 ($\tau=0.38$; $\alpha=\text{max}$. P–Pd–Cl=149°, $\beta=\text{max}$. P–Pd–P=172°). NBO charges: Cl _{ax} –0.872, Cl _{eq} –0.682.	
^a Structural index parameter τ as defined by Addison et al., ⁶⁰ for perfect square pyramidal geometry $\tau = 0$ and for perfect trigonal bipyramidal geometry $\tau = 1$.				

^a Structural index parameter τ as defined by Addison et al.,⁶⁰ for perfect square pyramidal geometry $\tau = 0$ and for perfect trigonal bipyramidal geometry $\tau = 1$.

A range of transition-metal complexes of the tetramethylbiphosphinine (tmbp, shown in Table 4a) ligand has been reported.⁵⁵ The square planar coordination geometry adopted by the ruthenium complex $[\text{Ru}(\text{tmbp})_2]^{2-}$ **14** is unusual, because a survey of tetracoordinate ruthenium (in any oxidation state, Table 2a) identified only three examples of this coordination geometry. In addition, if the tmbp ligand is considered as a neutral ligand, the metal in **14** can be assigned a d^{10} configuration for which a tetrahedral geometry would be expected, as confirmed by the DFT optimized geometry of dianionic **14** (Table 4b). Imposing D_{2h} symmetry, the square planar geometry with a closed-shell electronic configuration is only slightly higher in energy than the tetrahedral minimum (Table 4b). In the original structure report⁵⁴ the authors suggested that the tmbp ligands coordinate as radical anions, thus justifying the observed square planar geometry as being expected for a d^8 metal center. We have attempted to reproduce this electronic structure using broken-symmetry (BS)⁵⁶ DFT calculations on complex **14a**, where the methyl substituents have been removed from the tmbp ligand, but these radical dianion configurations lie more than 40 kcal mol^{-1} higher in energy than the tetrahedral energy minimum and the ligands are not planar (Table 4b). Imposing D_{2h} symmetry, planarity can be enforced, but the planar BS complex is energetically very unfavorable (Table 4b).

In a more recent publication Le Floch and co-workers⁵⁵ suggested that the electronic preference is indeed weakly in

favor of the tetrahedral geometry for $[\text{Ru}(\text{bp})_2]^{2-}$ **14a** but that the energy difference between tetrahedral and square planar geometry is small and no barrier to a flattening of the tetrahedron exists (summarized in Table 4a). However, in the crystal structure of **14** two $[\text{Li}(\text{thf})_3]^+$ counterions are weakly coordinated to the ruthenium. Le Floch et al. have treated these ions computationally using sodium cations, in a $[\text{Ru}(\text{bp})_2]\cdot\text{Na}_2$ complex, which was found to favor a square planar geometry. This suggests that despite the long $\text{Ru}\cdots\text{Li}$ distances (Table 4a) observed in the original crystal structure, the counterions are important for determining the observed geometry. Our own optimization of $[\text{Ru}(\text{tmbp})_2]\cdot[\text{Li}(\text{thf})_3]_2$ **14b** shows a slight twisting of the tmbp ligands away from square planar geometry (Table 4b, $\text{tmbp}\cdots\text{tmbp}$ dihedral = 21°), yet this geometry is in better agreement with the crystal structure molecular geometry than the isolated complex dianion **14**.

The pentacoordinate palladium(II) complex $\text{PdCl}_2(\text{PMe}_2\text{Ph})_3$ **15** has one long $\text{Pd}-\text{Cl}$ bond and one significantly shorter $\text{Pd}-\text{Cl}$ bond (Table 5b); the latter falls within the range normally observed for $\text{Pd(II)}-\text{Cl}$ fragments (Table 5a), while the long bond is clearly an outlier (0.626 \AA greater than the mean (2.330 \AA), equivalent to 12 SSDs). If the geometry is described as a distorted square pyramidal geometry, the longer $\text{Pd}-\text{Cl}$ bond is axial. The authors recognized the long $\text{Pd}-\text{Cl}$ distance as unusual and discussed whether the structure could be better described as a tight

ion pair, $[\text{PdCl}(\text{PMe}_2\text{Ph})_3]^+\cdot\text{Cl}^-$, but concluded that penta-coordination is more appropriate (summarized in Table 5b). This complex is also included in a structural correlation study of Aullón and Alvarez⁶¹ who discussed the structural changes associated with adduct formation (i.e. $[\text{ML}_4] + \text{L}' \rightarrow [\text{ML}_4\text{L}']$).

The DFT optimized geometry of **15** is closer to trigonal bipyramidal, with both chlorides in the equatorial plane (Table 5c). The longer Pd–Cl bond has contracted, whereas an elongation of the shorter bond can be observed. While the energy difference between trigonal bipyramidal and square-pyramidal coordination geometries is often not very large in 5-coordinate complexes,^{61,62} the crystallographically observed bond length difference is not confirmed by these calculations. To establish the influence of the more loosely coordinated chloride ion on the complex geometry, we have investigated complex $[\text{PdCl}(\text{PMe}_2\text{Ph})_3]^+$ **15a**. This complex optimizes to an almost perfectly square planar geometry with a short Pd–Cl bond (Table 5c), suggesting that the axial Pd–Cl interaction is important in determining the coordination geometry. This is in good agreement with the pyramidalization on adduct formation described by Alvarez.⁶¹ We have also investigated close contacts in the crystal structure environment and identified a reasonably short $\text{Cl}\cdots\text{HC}$ hydrogen bond interaction (2.564 Å) between the axial chloride and a dichloromethane molecule present in the crystal structure. DFT optimization of the adduct $[\text{PdCl}_2(\text{PMe}_2\text{Ph})_3]\cdot\text{CH}_2\text{Cl}_2$ **15b** gives a slightly longer Pd–Cl_{ax} bond and a shorter $\text{Cl}\cdots\text{H}$ interaction (2.309 Å), but agreement with the crystallographically observed geometry is still quite poor (Table 5c). The partial negative charge on the more remote axial chloride ligand might be stabilized by interactions with the surrounding crystal environment, but optimization in vacuo only allows a transfer of the charge density toward the metal center, presumably resulting in the observed discrepancies in Pd–Cl distances for **15** and **15b** (Table 5c). To test this hypothesis we investigated simulation of the environment as a continuum dielectric medium (CDM) as a simple and convenient way of mimicking this environmental/“solvation” effect for complexes with small anionic ligands (see for example ref 19 and references therein). We have therefore used a continuum dielectric medium as implemented in Jaguar to represent water solvation (described in Computational Details, Appendix). [It should be noted that the main structural change occurs on going from optimization in vacuo to a CDM, whereas changes in the dielectric constant of the “solvent” have only minor effects on M–L distances. Water solvation is implemented in most software packages and was chosen here for convenience.] The optimized geometry of solvated **15_{aq}** is in much better agreement with the observed crystal structure molecular geometry, reproducing the long axial Pd–Cl bond (Table 5c). In addition, the square-pyramidal geometry is reproduced slightly better than for the geometry optimized in vacuo. The continuum dielectric field stabilizes a larger partial charge on the axial chloride ligand, whereas for the complex optimized in vacuo the concentration of charge density on this chloride would be unfavorable and a shorter Pd–Cl distance thus reduces the polarity of this interaction. This is confirmed by the NBO charges calculated for the chloride ligands (Table 5c). While it may be important to represent the crystal environment if weak interactions influence the

observed geometries (see also discussion in part a, 2.i)), most crystal structure molecular geometries are determined by much stronger intramolecular steric and electronic interactions, and the optimization of isolated molecular units is usually sufficient.

The outliers discussed so far have either been directly confirmed by the computational approach, or the identification of a problem with the automation or computational model has led to a revised calculation giving a result in better agreement with the crystallographically observed molecular structure. However, the crystal structures deposited in the database may also contain errors. A number of potential sources of errors/mistakes during the structure refinement have been discussed in part a, but their conclusive identification can be difficult, especially if the original diffraction data are not available (vide supra). However, we have found that the geometry of the metal coordination sphere can be robust to refinements based on the wrong assignment of atomic identities (part a, 3.iv), i.e., bond lengths and angles agree with trends observed for the correct metal. Tables 6–8 illustrate three examples where a combination of database mining and computational experiments can be used to suggest alternative metal assignments. [It is noted that for the three structures discussed here with likely metal misassignment the *PLATON* validation software found no space group discrepancies or other major errors. This is not surprising as it appears that the wrong metal in all three structures is only evident from analysis of the metal coordination geometry, and, as Spek discussed with reference to *PLATON*, there is scope for the implementation of additional tests, particularly for inorganic structures.²⁷] In general, the identification of unusual M–L distances not confirmed by DFT calculations was followed by a search of the CSD for metal complexes of relevant ligands. This process was usually guided by the data in ref 13, but searches were repeated to take account of new crystal structure data. By considering complexes with similar ligands, metal oxidation state/complex charge, and favored coordination geometries, one or two likely alternative metal candidates could be identified. For these, DFT calculations were run to confirm improved agreement with the outlier crystal structure geometry.

A CSD survey of Cr–N bond lengths for primary amines and derivatives in 6-coordinate Cr(III) complexes (Table 6a) identified very short Cr–N distances (>5 SSDs from the mean value) for the $\text{NH}_4[\text{Cr}(\text{en})_2\text{ox}]\text{Cl}_2\cdot\text{H}_2\text{O}$ complex **18** (Table 6b). The DFT optimized structure of the cation $[\text{Cr}(\text{en})_2\text{ox}]^+$ **18a** does not reproduce the crystallographically observed short Cr–N distances (Table 6c), even when allowing for the DFT overestimation of M–L distances (discrepancy <0.08 Å vs mean deviation of 0.237 Å for **18a**). Another salt of $[\text{Cr}(\text{en})_2\text{ox}]^+$ has been reported in the CSD (**19**, Table 6b) in which the Cr–N bond lengths are in better agreement with both the Cr–N sample mean and the calculated geometry. A CSD search of M–N distances found good agreement between the crystal structure molecular geometry of **18** and the observed mean for cobalt(III) complexes, further confirmed by available data for the $[\text{Co}(\text{en})_2\text{ox}]^+$ cation (Table 6a) and its DFT optimized geometry (**20**, Table 6c). Since both the DFT results and the CSD survey sample mean for the related *cobalt* complex are in much better agreement with the structural parameters of complex **18**, and the authors have actually investigated

Table 6. Summary of Results Related to a Survey of Cr–N Bond Lengths (Primary Amines and Derivatives) in Hexacoordinate Cr(III) Complexes

(a) Summary of Database Search						
CSD search fragment			no. of fragments		summary of results	
6-coordinate Cr(III)-NH ₂ C, (N–C cyclic bond)			329		mean Cr–N = 2.079(23) Å	
6-coordinate Cr(III)-oxalate (chelating (O ₂ C) ₂)			166		mean Cr–O = 1.963(18) Å	
6-coordinate Co(III)-NH ₂ C, (N–C cyclic bond)			2765		mean Co–N = 1.964(22) Å	
[Co(III)(en) ₂ ox] ⁺ cation ^a			8		Co–N range: 1.924–1.988 Å; Co–O range: 1.883–1.936 Å	
(b) Description of Crystal Structures						
no.	CSD Refcode	complex	M–N/Å	M–O/Å	summary of structure report	ref
18	POFHEX	NH ₄ [Cr(en) ₂ ox]Cl ₂ ·H ₂ O	1.921–1.944	1.910, 1.921	No comment on short Cr–N. Authors compared the crystallization behavior of 18 with related cobalt complexes [Co(en) ₂ ox]X (X = halide)	63
19	CRENOX	[Cr(en) ₂ ox][Cr(en)(ox) ₂]·2H ₂ O	2.064–2.072	1.958, 1.958		64
(c) Calculation Results						
no.	input description			M–N/Å	M–O/Å	
18a	[Cr(en) ₂ ox] ⁺ , from XRD			2.144–2.196	1.885, 1.885	
20	[Co(en) ₂ ox] ⁺			1.987–2.036	1.859, 1.860	
^a en = 1,2-ethylenediamine- <i>N,N</i> ; ox = oxalato- <i>O,O</i> .						

^a en = 1,2-ethylenediamine-*N,N*; ox = oxalato-*O,O*.**Table 7.** Summary of Results for Tetracoordinate Palladium Complexes

(a) Summary of Database Search						
CSD search fragment		no. of fragments	summary of results			
4-coord. Pd		>4000	Predominantly square planar, <20 tetrahedral complexes, but all Pd(0), d ¹⁰ . Of group 10 d ⁸ metals, only high spin (<i>S</i> = 1) Ni(II) complexes tetrahedral.			
4-coord. Pd(II)-Cl		1200	mean Pd–Cl = 2.328(40) Å, ³ all square planar, minimum Pd–Cl = 2.250 Å {DICSIR Pd–Cl bond lengths (2.222, 2.204 Å) (Table 7b)}			
4-coord. Ni(II)-Cl		167	mean Ni–Cl: 2.218(36) Å, square planar and tetrahedral			
(b) Description of Outlier						
no.	CSD Refcode	complex	A1, A2/°	Pd–Cl/Å	summary of structure report	ref
21	DICSIR	[PdCl ₂ (dppfO ₂)] ^a	120, 104	2.222, 2.204.	described as a tetrahedral, paramagnetic Pd complex.	65
(c) Calculation Results						
no.	input description		A1, A2/°	M-Cl/Å	rel <i>E</i> /kcal mol ^{−1}	
21	from XRD, <i>S</i> = 0		177, 177	2.311, 2.310	0.0	
	from XRD, <i>S</i> = 1		145, 94	2.427, 2.369	20.5	
22	[NiCl ₂ (dppfO ₂)], <i>S</i> = 0		177, 175	2.198, 2.205	22.0	
	<i>S</i> = 1		137, 100	2.256, 2.289	0.0	
^a dppfO ₂ = 1,1′-bis(oxodiphenylphosphoranyl)ferrocene.						

^a dppfO₂ = 1,1'-bis(oxodiphenylphosphoranyl)ferrocene.

the cobalt complex in the past, this suggests that the “chromium” assignment in the reported structure of **18** (CSD Refcode POFHEX) may be in error and that the crystal structure refinement was based on the wrong transition metal. It thus seems likely that POFHEX is the crystal structure of NH₄[Co(en)₂ox]Cl₂·H₂O. [While the colors of Cr(III) and Co(III) oxalate complexes are strikingly different (reddish violet and dark green respectively), their mixed oxalate/ethylenediamine complexes discussed here are less distinctive (described as orange and red respectively). In more general terms, we do not consider the crystal color descriptions in the CSD as reliable indicators of metal identity and electronic configuration, as these are often not given, or can be subjective and dependent on crystal size and shape as well as varying with counterions. Consideration of UV/vis absorption maxima (λ_{max}) could provide some indication of the metal ion identity, but these data are often not available.]

Tetracoordinate palladium(II) complexes are square planar (see Table 7a); crystal structures of tetrahedral palladium complexes have been reported, but they all correspond to Pd(0), where the tetrahedral geometry is favored for the d¹⁰ configuration. In their survey of d⁸ palladium(II) and platinum(II) tetracoordinate complexes, Alvarez and co-workers¹⁴ did not observe any tetrahedral coordination geometries. Of the group 10 d⁸ metals, only nickel complexes with a high spin (*S* = 1) configuration have tetrahedral geometry. However, one crystal structure with tetrahedral, paramagnetic palladium(II) has been reported, that of [PdCl₂(dppfO₂)] **21**. One of the phenyl substituents of the dppfO₂ ligand is disordered, so it would have been excluded from any of our database searches. However, since a tetrahedral geometry is very unusual for a palladium(II) complex and the disorder should not affect the metal–ligand coordination, we have investigated this complex in detail.

Table 8. Summary of Results Relevant to *cis*-[RuCl₂(C≡N(2,6-Me₂-Ph))₂] Complex **23**

(a) Summary of Database Search (See also Table 2a)						
CSD search fragment			no. of fragments	summary of results		
Ru(II)–Cl			1416	mean Ru–Cl = 2.421(32) Å		
Ru(II)–CNR (terminal isocyanide, R = any C)			86	mean Ru–C = 1.979(47) Å		
4-coordinate Pd(II)–Cl, with sp ¹ carbon trans to Cl			22	mean Pd–Cl ³ = 2.307(25) Å		
4-coordinate Pd(II)–CNR (terminal isocyanide) with Cl <i>trans</i> to isocyanide			18	mean Pd–C = 1.924(25) Å		
(b) Description of Complexes						
no.	CSD Refcode	complex	A1, A2/°	M–Cl/Å	M–C ^a /Å	ref
23	YIMLEL	<i>cis</i> -[RuCl ₂ (C≡N–(2,6-Me ₂ -Ph)) ₂]	178, 177	2.311, 2.307	1.925, 1.937	67
24	SIVGAF	<i>cis</i> -[PdCl ₂ (C≡N–(2,6-Me ₂ -Ph)) ₂]	178, 177	2.303, 2.294	1.906, 1.915	68
(c) Calculation Results						
no.	input description	A1, A2/°	M–Cl/Å	M–C ^a /Å	comments	
23	from XRD, <i>S</i> = 0	156, 96	2.363, 2.364	1.869, 1.869	rel <i>E</i> = 0.0 kcal mol ^{–1}	
	from XRD, <i>S</i> = 1	168, 165	2.370, 2.374	1.991, 1.998	rel <i>E</i> = 14.9 kcalmol ^{–1}	
24	from XRD, <i>S</i> = 0	177, 177	2.352, 2.352	1.964, 1.964		
^a Isocyanide.						

^a Isocyanide.

We have optimized **21** with both low and high spin electronic configurations, starting from the crystal structure molecular geometry. The singlet optimizes to a square planar complex with Pd–Cl bond lengths in reasonably good agreement with the CSD sample mean, whereas the paramagnetic triplet configuration suggested by the experimentally determined magnetic moment gives a distorted complex higher in energy (Table 7c). Since the sample mean for Ni(II)–Cl bonds (Table 7a) is in much closer agreement with the observed structure, we have also investigated the corresponding nickel complex **22** in both electronic configurations (Table 7c). In this case the singlet configuration again gives rise to a square planar complex but is energetically disfavored, whereas the optimized high spin complex has an approximately tetrahedral geometry with Ni–Cl distances in reasonably good agreement with the CSD sample mean (Table 7a, 2.218(36) Å). The authors discuss the preparation of the nickel complex **22** in the structure report; again, it seems, the molecular formula used in the crystal structure refinement may have been in error. The experimental diffraction data are indeed fitted better when nickel is used instead of palladium in the structure refinement.⁶⁶

As discussed (*vide supra*), the favorable coordination geometry for tetracoordinate ruthenium(II) complexes appears to be a disphenoidal ligand arrangement, stabilized by both electronic and steric/agostic interactions. However, tetracoordinate ruthenium complexes are rare in the CSD (Table 2a), and a general statement cannot be derived from the available crystal structure data alone. We have therefore investigated the square planar complex *cis*-[RuCl₂(C≡N(2,6-Me₂-Ph))₂] **23** (Table 8).

The Ru(II)–Cl bond lengths of **23** are the shortest found in the CSD survey (Table 8a) and are outliers (> 3 SSDs from the mean), with the Ru(II)–C bond lengths also shorter than the database mean. [It should be borne in mind, though, that the CSD values are based on all coordination geometries, not just tetracoordinate complexes.] DFT optimization of **23** gave a disphenoidal complex with Ru–Cl bond lengths in reasonable agreement with the crystal structure molecular geometry but underestimated the isocyanide Ru–C distances (Table 8c). One other example of a ruthenium(II) complex adopting a square planar geometry in the crystal structure

has been reported (CSD Refcode OKISUW, see Table S1).⁶⁹ In this case the tridentate ligand PNP ligand enforces a square planar geometry, but the complex is paramagnetic and the reported DFT calculations confirmed that the open-shell triplet (*S* = 1) electronic configuration is the ground state. We have therefore optimized **23** with a triplet configuration (*S* = 1, intermediate spin). This complex lies 14.9 kcal mol^{–1} higher in energy than the singlet, but an approximately square planar geometry with bond lengths in better agreement with the crystal structure is adopted (Table 8c).

Few square planar transition-metal complexes with at least two isocyanide ligands have been reported in the CSD, and with the exception of the ruthenium(II) complex **23**, these are either palladium(II) or rhodium(I) complexes, where the d⁸ configuration is expected to favor a square planar geometry. However, since the corresponding rhodium(I) complex has a different charge, we have only considered palladium(II) here. The crystallographic mean Pd–Cl and Pd–CNR distances with Csp¹ or Cl trans respectively are in good agreement with the observed data for **23** (Table 8a,b). One palladium(II) structure in the CSD (SIVGAF, **24**) has the same ligands as the ruthenium(II) complex **23**, and it adopts a square planar geometry with Pd–Cl and Pd–C bond lengths similar to those observed for Ru–Cl and Ru–C in **23** (complex **24**, Table 8b). The DFT optimized palladium complex **24** (Table 8c) maintains the expected square planar geometry, and the calculated Pd–Cl bond lengths are in reasonably good agreement with both the CSD sample mean value and the M–Cl distances observed in the crystal structures of **23** and **24** (Table 8a,b). Both calculated coordination geometry and metal–ligand distances are in better agreement for the palladium complex **24** than for the equivalent ruthenium(II) complex **23**, suggesting again that the assignment of the metal as ruthenium in the crystal structure of **23** is in error. In this case, further tentative evidence arises from comparison of the space group and reduced cell parameters reported in the CSD for **23** and **24**. Both structures have the same space group (P-1) and the reduced cell parameters are very similar, although with slight deviations, particularly for the *b* axis length. The unit cell parameters and the slightly shorter metal–ligand bond lengths observed for **24** (SIVGAF, Pd) suggest that the data

collection was carried out at a lower temperature than for **23** (YIMLEL, refined as Ru), although both are apparently room-temperature structure determinations.

SUMMARY AND OUTLOOK

A protocol for the addition of knowledge value to crystal structure databases and their derived knowledge bases using computational chemistry has been described. This protocol is aimed at verifying transition-metal complex geometries detected as outliers or unusual by a knowledge base generation software such as (*Metallo-organic*) *Mogul* and is intended to take advantage of developments in electronic archiving, database access, and distributed computing emerging from the field of e-science.

The first stage of this procedure is envisaged to be fully automated and consists of outlier identification, followed by DFT optimization and comparison with the observed crystal structure molecular geometry to distinguish between structures confirmed by calculation and those that are not. Irrespective of the result of this comparison, the next stage requires inspection of the original structure report and related publications to determine whether the unusual structural feature has been recognized and explained. The database (or knowledge base) entry can then be augmented with a summary of this explanation as well as relevant experimental data. If the explanation is incomplete or missing, the calculated results can complement the structure report and should thus be added to the database. The scope for extracting “knowledge value” from the calculated results, e.g. by visualizing molecular orbitals and by comparing energies for different coordination geometries and/or electronic configurations, has been illustrated with a number of representative examples. In cases where the crystallographically observed geometry is not confirmed by the DFT calculation, a number of potential problems/weaknesses arising from the chosen computational approach need to be investigated, and the most likely issues have been illustrated in this work. The discrepancy can also arise from an erroneous crystal structure, and a range of problems has been identified. Often these crystallographic problems cannot be proven, unless the original experimental data is available for re-refinement. However, we have discussed several examples of the structurally most robust issue, i.e. where a wrong metal assignment may have been used in the refinement, and illustrated how a combination of database mining and computational chemistry can be used to guide alternative solutions.

While the automation of the first stage of this procedure has been discussed explicitly, we feel that later stages, such as the literature searching and results analysis/expansion, can be achieved more *efficiently* with some human intervention, because these processes rely on an understanding of the chemical context of the structural results. This preference reflects our current valuation of the “costs” of calculations versus that of user involvement for implementation and analysis. In brief, with current resources DFT calculations on large transition-metal complexes cost considerable computational time, and it seems desirable both to avoid unnecessary calculations and to maximize the utility of results by appropriate analysis and archiving. The time of programmers is also valuable and, at least initially, should be focused on implementing automated protocols for stan-

dard, repetitive procedures as in stage 1. While at least some of the additional calculations in stage 2 could potentially be automated, the chemical expertise implicit in choosing the *most relevant* approach for further investigation would be very difficult to capture computationally. Computational costs require that calculations are relevant; in addition, researcher time should not be wasted on the analysis of irrelevant results. We therefore suggest a partially automated approach for stage 2, where researcher intervention is focused on establishing the chemical context, choosing from a range of automated and manual calculation options, and analyzing and presenting results, thus maximizing the benefits of researcher expertise.

However, future e-science developments are likely to change both the costs of calculations and of the archiving and exploitation of data. While this will mainly speed up knowledge base generation in stage 1, it will also facilitate a more extensive automation of stage 2 procedures:

(i) *Chemical Context.* Electronic archiving of published research and the underlying experimental data can facilitate the automation of literature searches if information is organized in a standardized way and marked up with appropriate metadata tags.⁷⁰ This will allow automatic checking for additional experimental data when a crystal structure molecular geometry has been identified as an outlier and hence to initiate further calculations based on, for example, different electronic configurations. In addition, summary explanations, additional calculation results, and protocols of user intervention can be archived in context with the original structure report and related work.

(ii) *Implementation.* Another useful aspect of these e-science developments will be that the more widespread availability of high performance computing resources will reduce the need for efficient data generation and analysis. Rather than manually choosing the most likely explanation before initiating additional calculations, intervention processes can be recorded/observed once and then replayed by an automatic system whenever a similar situation is encountered. This will mainly lead to the automation of calculations aimed at exploring different electronic configurations and starting geometries as well as the explicit (counterions, periodic images of unit cell) or implicit (CDM) consideration of the crystal environment. More “unusual”/difficult procedures (BS calculations, computational experiments) could also be suggested automatically but would continue to rely on manual confirmation and user intervention. With appropriate feedback loops this initially very inefficient approach could gradually be turned into a chemical expert system for detection, optimization, and analysis of unusual molecular geometries observed in crystal structures, thus further focusing the user involvement on the final knowledge result rather than the identification of computational and crystallographic problems.

(iii) *Analysis.* The detailed analysis and presentation of results (e.g. generation of orbital energy level diagrams, extraction of molecular orbital pictures, analysis of charge distributions) could also be recorded and automated, thus further reducing the demands on researcher time.

While the field of e-science is still in its infancy, the design of a knowledge base as reported in this work or our prototype ligand knowledge base¹ is aimed at accommodating these developments by formalizing the underlying protocols.

ACKNOWLEDGMENT

The authors would like to thank J. P. H. Charmant and M. F. Haddow for useful discussions of crystallography and R. Taylor, F. H. Allen, and other staff of the Cambridge Crystallographic Data Center (CCDC) for helpful discussions of knowledge base development and for providing access to the Cambridge Structural Database (CSD). Financial support of the Engineering and Physical Sciences Research Council (EPSRC), including an Advanced Research Fellowship (J.N.H.), is gratefully acknowledged.

APPENDIX

Database Mining Details. (a) *Tetracoordinate Geometries.* Four-coordinate metal complexes of V, Cr, W, Re, Ru, Ni, and Pd were retrieved from the CSD (Nov 2003). The coordination at the metal was defined by the spatial coordination number (SCN)¹³ which takes into account the number of coordination sites occupied by multihapto and bridging ligands rather than simply counting the total number of ligand atoms coordinated to the metal. Therefore η^2 -ligands occupy one coordination site only, and for fragments with a M–M' bond bridged by one ligand atom the M' atom is not counted. Metal cluster fragments where the central metal is coordinated to two or more other metals were excluded. All entries with $R < 0.10$ with no errors or disorder were included in the geometry searches. Duplicate structures were removed.

As described previously,³ the 4-coordinate geometry was defined simply by the maximum L^1-M-L^2 angle and the angle defined by the other 2 ligands, L^3-M-L^4 . For η^2 -ligands the centroid of the L–L bond was taken to define the angle. The standard geometries for 4-coordinate species were then defined as follows:

square planar: $L^1-M-L^2, L^3-M-L^4 > 150^\circ$

tetrahedral: $L^1-M-L^2, L^3-M-L^4 < 140^\circ$

disphenoidal: $L^1-M-L^2 > 150^\circ; L^3-M-L^4 < 130^\circ$

All other possible geometries were defined simply as 'distorted' geometries. The maximum angle L^1-M-L^2 is labeled A1 and the L^3-M-L^4 angle A2. For each metal the distribution of geometries using the angle descriptors A1 and A2 was analyzed, and a selection of structures with unusual geometries was identified and investigated computationally. We note that a more detailed analysis of 4-coordinate geometry could be achieved, for example, by using symmetry measures as employed by Alvarez.¹⁴

(b) *Unusual M–L Bond Lengths.* Based on searches of the CSD for various metal–ligand fragments, M–L bond length outliers were identified. Outliers were defined as deviating by more than 3 sample standard deviations (SSDs) from the mean bond length observed for a given fragment. The outliers discussed here were obtained from analysis of data sets for metal–ligand fragments containing terminal chloride, monodentate phosphorus(III), or primary amine ligands. The data sets were obtained from the CSD, and only structures which were described as having no disorder or unexplained errors and with R -factor less than 0.075 were included.

Computational Details. All calculations used the Jaguar package^{71,72} and the standard B3LYP density functional.⁷³ Open shell electronic configurations used the Jaguar default restricted-open shell (RODFT) formalism; in some cases spin unrestricted density functional theory (UDFT) was also used. The Jaguar triple- ζ form of the standard Los Alamos ECP basis set (LACV3P) was used on V, Cr, W, Re, Fe, Ru, Ni, Pd, Br, and I, with the 6-31G* basis used for all other atoms. In the few cases where calculations were slow to converge, "loose" convergence (5 times larger than default criteria) was used for geometry optimizations. Test calculations using the more stringent default convergence criteria did not lead to significant changes in energies, bond lengths, or angles but were much more time-consuming. Broken symmetry calculations⁵⁶ were performed for complex **14a**. Water solvation was simulated for complex **15_{aq}** using a self-consistent reaction field dielectric continuum⁷⁴ as implemented in Jaguar 5.0⁷¹ with default settings for water. Unless stated otherwise, calculations were performed on isolated molecules and NBO atomic charges⁷⁵ were calculated.

Supporting Information Available: Table of complexes where structural feature explained in detail in original structure report (S1) and tables summarizing main structural features (M–L, A1, A2, relevant others) for DFT optimized geometries (S2). This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES AND NOTES

- (1) Fey, N.; Tsipis, A.; Harris, S. E.; Harvey, J. N.; Orpen, A. G.; Mansson, R. A., Development of a Ligand Knowledge Base. Computational Descriptors for Phosphorus Donor Ligands. *Chem. Eur. J.* **2006**, *12*, 291–302.
- (2) Bruno, I. J.; Cole, J. C.; Kessler, M.; Luo, J.; Motherwell, W. D. S.; Purkis, L. H.; Smith, B. R.; Taylor, R.; Cooper, R. I.; Harris, S. E.; Orpen, A. G. Retrieval of Crystallographically-Derived Molecular Geometry Information. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2133–2144.
- (3) Harris, S. E.; Orpen, A. G.; Bruno, I. J.; Taylor, R., Factors Affecting d-Block Metal–Ligand Bond Lengths: Towards An Automated Library of Molecular Geometry For Metal Complexes. *J. Chem. Inf. Model.* **2005**, *45*, 1727–1748.
- (4) Allen, F. H. The CSD: a quarter of a million structures and rising. *Acta Crystallogr.* **2002**, *B58*, 380–388.
- (5) Orpen, A. G. Applications of the Cambridge Structural Database to molecular inorganic chemistry. *Acta Crystallogr.* **2002**, *B58*, 398–406, and reference cited.
- (6) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Feng, Z.; Gilliland, G.; Iype, L.; Jain, S.; Fagan, P.; Marvin, J.; Padilla, D.; Varichandran, V.; Schneider, B.; Thanki, N.; Weissig, H.; Westbrook, J. D.; Zardecki, C., The Protein Data Bank. *Acta Crystallogr.* **2002**, *D58*, 899–907. RSCB, Protein Data Bank. <http://www.pdb.org/> (22/02/2005).
- (7) Crystallographic databases. Joint special issue with *Acta Crystallographica Section D*. *Acta Crystallogr.* **2002**, *B58*, 317–422, and references cited.
- (8) Crystallographic databases. Joint special issue with *Acta Crystallographica Section B*. *Acta Crystallogr.* **2002**, *D58*, 879–920, –920, and references cited.
- (9) Allen, F. H.; Taylor, R. Research applications of the Cambridge Structural Database (CSD). *Chem. Soc. Rev.* **2004**, *33*, 463–475, and references cited.
- (10) Bruno, I. J.; Cole, J. C.; Lommerse, J. P. M.; Rowland, R. S.; Taylor, R.; Verdunk, M., IsoStar: A library of information about nonbonded interactions. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 525–537.
- (11) Defining e-Science. <http://www.nesc.ac.uk/nesc/define.html> (04/06/2004).
- (12) Allen, F. H.; Kennard, O.; Watson, D. G.; Brammer, L.; Orpen, A. G.; Taylor, R. Tables of bond lengths determined by X-ray and neutron diffraction. Part 1. Bond lengths in organic compounds. *J. Chem. Soc., Perkin Trans. 2* **1987**, S1–S19.

- (13) Orpen, A. G.; Brammer, L.; Allen, F. H.; Kennard, O.; Watson, D. G.; Taylor, R., Supplement. Tables of bond lengths determined by X-ray and neutron diffraction. Part 2. Organometallic compounds and coordination complexes of the d- and f-block metals. *J. Chem. Soc., Dalton Trans.* **1989**, S1–S83.
- (14) Cirera, J.; Alemany, P.; Alvarez, S., Mapping the Stereochemistry and Symmetry of Tetracoordinate Transition-Metal Complexes. *Chem. Eur. J.* **2004**, *10*, 190–207, and references cited.
- (15) See, R. F.; Kruse, R. A.; Strub, W. M. Metal–Ligand Bond Distances in First-Row Transition Metal Coordination Compounds: Coordination Number, Oxidation State, and Specific Ligand Effects. *Inorg. Chem.* **1998**, *37*, 5369–5375. Rulíxf0ek, L.; Vondráxf0ek, J. Coordination geometries of selected transition metal ions in metalloproteins. *J. Inorg. Biochem.* **1998**, *71*, 115–127.
- (16) Bray, M. R.; Deeth, R. J.; Paget, V. J.; Sheen, P. D. The Relative Performance of the Local Density Approximation and Gradient-Corrected Density Functional Theory for Computing Metal–Ligand Distances in Werner-Type and Organometallic Complexes. *Int. J. Quantum Chem.* **1996**, *61*, 85–91.
- (17) Matveev, A.; Staufer, M.; Mayer, M.; Roesch, N., Density Functional Study of Small Molecules and Transition-Metal Carbonyls Using Revised PBE Functionals. *Int. J. Quantum Chem.* **1999**, *75*, 863–873.
- (18) Salomon, O.; Reiher, M.; Hess, B. A. Assertion and validation of the performance of the B3LYP* functional for the first transition metal row and the G2 test set. *J. Chem. Phys.* **2002**, *117*, 4729–4737.
- (19) Deeth, R. J.; Fey, N. The performance of nonhybrid density functionals for calculating the structures and spin states of Fe(II) and Fe(III) complexes. *J. Comput. Chem.* **2004**, *25*, 1840–1848, and references cited.
- (20) Norrby, P.-O.; Brandt, P. Deriving force field parameters for coordination complexes. *Coord. Chem. Rev.* **2001**, *212*, 79–109. Hocking, R. K.; Hambley, T. W. Statistical and molecular mechanics analysis of the effects of changing donor type on bond length in the two series [(CoN_nO_{6-n})–N–III] and [(NiN_nO_{6-n})–N–II] (*n* = 0–6): A new route to bond-stretch parameters. *Inorg. Chem.* **2002**, *41*, 2660–2666. David, L.; Amara, P.; Field, M. J.; Major, F. Parametrization of a force field for metals complexed to biomacromolecules: applications to Fe(II), Cu(II) and Pb(II). *J. Comput.-Aided Mol. Des.* **2002**, *16*, 635–651. Marques, H. M.; Cukrowski, I. Molecular mechanics modelling of porphyrins. Using artificial neural networks to develop metal parameters for four-coordinate metallocporphyrins. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5878–5887.
- (21) Martín, A.; Orpen, A. G. Structural Systematics. 6. Apparent Flexibility of Metal Complexes in Crystals. *J. Am. Chem. Soc.* **1996**, *118*, 1464–1470.
- (22) Computational Transition Metal Chemistry. *Chem. Rev.* **2000**, *100*, 351–818, and references cited.
- (23) Allen, F. H.; Harris, S. E.; Taylor, R. Comparison of conformer distributions in the crystalline state with conformational energies calculated by ab initio techniques. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 247–254.
- (24) Reiher, M.; Salomon, O.; Hess, B. A. Reparametrization of hybrid functionals based on energy differences of states of different multiplicity. *Theor. Chem. Acc.* **2001**, *107*, 45–55. Swart, M.; Groenhof, A. R.; Ehlers, A. W.; Lammersma, K. Validation of Exchange–Correlation Functionals for Spin States of Iron Complexes. *J. Phys. Chem. A* **2004**, *108*, 5479–5483. Fouqueau, A.; Mer, S.; Casida, M. E.; Daku, L. M. L.; Hauser, A.; Mineva, T.; Neese, F. Comparison of density functionals for energy and structural differences between the high-⁵T_{2g}; (t_{2g})⁴(e_g)² and low-¹A_{1g}; (t_{2g})⁶(e_g)⁰ spin states of the hexaquoferrous cation [Fe(H₂O)₆]²⁺. *J. Chem. Phys.* **2004**, *120*, 9473–9486. Fouqueau, A.; Casida, M. E.; Daku, L. M. L.; Hauser, A.; Neese, F. Comparison of density functionals for energy and structural differences between the high-⁵T_{2g}; (t_{2g})⁴(e_g)² and low-¹A_{1g}; (t_{2g})⁶(e_g)⁰ spin states of iron(II) coordination compounds. II. More functionals and the hexaminoferrous cation, [Fe(NH₃)₆]²⁺. *J. Chem. Phys.* **2005**, *122*, 044110;1–13.
- (25) Harvey, J. N. DFT Computation of Relative Spin-State Energetics of Transition Metal Compounds. *Struct. Bonding* **2004**, *112*, 151–184.
- (26) Rappé, A. K.; Bernstein, E. R. Ab Initio Calculation of Nonbonded Interactions: Are We There Yet? *J. Phys. Chem. A* **2000**, *104*, 6117–6128. Grimme, S. Accurate Description of van der Waals Complexes by DFT including empirical corrections (DFT-D). *J. Comput. Chem.* **2004**, *25*, 1463–1473. Johnson, E. R.; Wolkow, R. A.; DiLabio, G. A. Application of 25 density functionals to dispersion-bound homonuclear dimers. *Chem. Phys. Lett.* **2004**, *394*, 334–338.
- (27) Spek, A. L. Single-crystal structure validation with the program PLATON. *J. Appl. Crystallogr.* **2003**, *36*, 7–13.
- (28) Marsh, R. E.; Spek, A. L. Use of software to search for higher symmetry: space group C2. *Acta Crystallogr.* **2001**, *B57*, 800–805. Clemente, D. A.; Marzotto, A. 22 Space-group changes. *Acta Crystallogr.* **2003**, *B59*, 43–50.
- (29) Marsh, R. E. Space Group Cc: an update. *Acta Crystallogr.* **2004**, *B60*, 252–253.
- (30) Spek, A. L. PLATON, An Integrated Tool for the Analysis of the Results of a Single-Crystal Structure Determination. *Acta Crystallogr.* **1990**, *A46*, C34.
- (31) Cundari, T. R.; Harvey, J. N.; Klinckman, T. R.; Fu, W. Multiple Bonding Involving Late Transition Metals. The Case of a Silver-Oxo Complex. *Inorg. Chem.* **1999**, *38*, 5611–5615. Cundari, T. R.; Harvey, J. N.; Klinckman, T. R.; Fu, W. Additions and Corrections. *Inorg. Chem.* **2000**, *39*, 1336.
- (32) Listemann, M. L.; Schrock, R. R.; Dewan, J. C.; Kolodziej, R. M. Synthesis and Reactivity of Two Monomeric Tungsten(IV) Phenoxide Complexes. *Inorg. Chem.* **1988**, *27*, 264–271.
- (33) Scott, M. J.; Wilish, W. C. A.; Armstrong, W. H. Unprecedented Example of Four Coordination at a Vanadium(II) Center. Synthesis, Structure and Properties of a Reactive, Nearly Planar V(II) Phenolate Complex, [V(DIPP)₄{Li(THF)}₂] (DIPP = 2,6-Diisopropylphenolate). *J. Am. Chem. Soc.* **1990**, *112*, 2430–2432, and references cited.
- (34) Minhas, R. K.; Edema, J. J. H.; Gambarotta, S.; Meetsma, A. Stability of Vanadium(II) Aryloxides: Synthesis and Characterization of Sterically Protected, Neutral and Monomeric Vanadium(II) Aryloxides. Reactivity with the N–N Bond of (Me₃Si)CH=N=N. *J. Am. Chem. Soc.* **1993**, *115*, 6710–6717.
- (35) Gardiner, I. M.; Bruck, M. A.; Wexler, P. A.; Wigley, D. E. Monomeric Rhenium (IV) Phenoxides and Their Development from Mononuclear Rhenium Halides with Oxygen and Sulfur Donor Ligands. *Inorg. Chem.* **1989**, *28*, 3688–3695.
- (36) Cayton, R. H.; Chisholm, M. H.; Clark, D. L.; Hammond, C. E. Mo(OH)₄ⁿ⁺ (n=0,2), Tetrahedral or Square-Planar? The dⁿ-d¹⁰⁻ⁿ Hole Formalism for π-Donor and π-Acceptor Ligand Fields. *J. Am. Chem. Soc.* **1989**, *111*, 2751–2755.
- (37) Durfee, L. D.; Latesky, S. L.; Rothwell, I. P.; Huffman, J. C.; Folting, K. Chemical and electrochemical reduction of titanium(IV) aryloxides. *Inorg. Chem.* **1985**, *24*, 4569–4573.
- (38) Elian, M.; Hoffmann, R. Bonding Capabilities of Transition Metal Carbonyl Fragments. *Inorg. Chem.* **1975**, *14*, 1058–1076. Burdett, J. K.; Fässler, T. F. Pair Potentials, Electron-Pair Repulsions, and the Geometries of Molecules. *Inorg. Chem.* **1991**, *30*, 2859–2868.
- (39) Baratta, W.; Herdtweck, E.; Rigo, P. [RuCl₂](PPh₂(2,6-Me₂C₆H₃))₂]: A neutral 14-electron Ru(II) complex with 2 agostic interactions. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 1629–1631.
- (40) Huang, D.; Streib, W. E.; Bollinger, J. C.; Caulton, K. G.; Winter, R. F.; Scheiring, T. 14-Electron Four-Coordinate Ru(II) Carbonyl Complexes and Their Five-Coordinate Precursors: Synthesis, Double Agostic Interactions, and Reactivity. *J. Am. Chem. Soc.* **1999**, *121*, 8087–8097.
- (41) Huang, D.; Folting, K.; Caulton, K. G. Silyl Migration of Me₃SiCCPh Coordinated to [RuH(CO)(P^tBu₂Me)₂][BAR'₄] Can Be Reversed: Synthesis and Structure of [Ru(CH=C(SiMe₃)(Ph))(CO)(P^tBu₂Me)₂][BAR'₄]. *J. Am. Chem. Soc.* **1999**, *121*, 10318–10322.
- (42) Catalá, R.-M.; Cruz-Garritz, D.; Sosa, P.; Terreros, P.; Torrens, H.; Hills, A.; Hughes, D. L.; Richards, R. L. Ru–X–C (X = F or H) interactions in complex compounds. Preparation and crystal structures of [Ru{SC₆F₄(F-2)}(SC₆F₅)₂(PMe₂Ph)₂] and [Ru(SC₆F₅)₂PC₆H₄(H-2)](C₆H₅)₂]. *J. Organomet. Chem.* **1989**, *359*, 219–232.
- (43) Poli, R. Open-Shell Organometallics as a Bridge between Werner-Type and Low-Valence Organometallic Complexes. The Effect of the Spin State on the Stability, Reactivity, and Structure. *Chem. Rev.* **1996**, *96*, 2135–2204.
- (44) Hawrelak, E. J.; Bernskoetter, W. H.; Lobkovsky, E.; Yee, G. T.; Bill, E.; Chirik, P. J. Square Planar vs Tetrahedral Geometry in Four Coordinate Iron(II) Complexes. *Inorg. Chem.* **2005**, *44*, 3103–3111.
- (45) Baratta, W.; Mealli, C.; Herdtweck, E.; Ienco, A.; Mason, S. A.; Rigo, P. Nonclassical vs Classical Metal···H₃C–C Interactions: Accurate Characterization of a 14-Electron Ruthenium(II) System by Neutron Diffraction, Database Analysis, Solution Dynamics and DFT Studies. *J. Am. Chem. Soc.* **2004**, *126*, 5549–5562.
- (46) Huang, D.; Huffman, J. C.; Bollinger, J. C.; Eisenstein, O.; Caulton, K. G., The first η²-CH₂Cl₂ Adduct of Ru(II): [RuH(η²-CH₂Cl₂)(CO)-(P^tBu₂Me)₂][BAR'₄] (Ar' = 3,5-C₆H₃(CF₃)₂) and Its RuH(CO)-(P^tBu₂Me)₂⁺ Precursor. *J. Am. Chem. Soc.* **1997**, *119*, 7398–7399. Cooper, A. C.; Streib, W. E.; Eisenstein, O.; Caulton, K. G. *tert*-butyl is superior to phenyl as an agostic donor to 14-electron Ir(III). *J. Am. Chem. Soc.* **1997**, *119*, 9069–9070.
- (47) Ujaque, G.; Cooper, A. C.; Maseras, F.; Eisenstein, O.; Caulton, K. G. Computational Evidence of the Importance of Substituent Bulk on Agostic Interactions in Ir(H)₂(P^tBu₂Ph)₂⁺. *J. Am. Chem. Soc.* **1998**, *120*, 361–365.
- (48) Burdett, J. K. Calculation of the Geometries of Binary Transition Metal Carbonyl and Dinitrogen Complexes. *J. Chem. Soc., Faraday Trans. 2* **1974**, 1599–1613.
- (49) Pearson, R. G. Concerning Jahn-Teller Effects. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 2104–2106.

- (50) Barratta, W.; Herdtweck, E.; Rigo, P. [RuCl₂{PPh₂(2,6-Me₂C₆H₃)₂}]₂: A neutral 14-electron Ru(II) complex with 2 agostic interactions. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 1629–1631.
- (51) Walker, J. D.; Poli, R. FeCl₃-Phosphine Adducts with Trigonal-Bipyramidal Geometry. Influence of the Phosphine on the Spin State. *Inorg. Chem.* **1989**, *28*, 1793–1801.
- (52) Godfrey, S. M.; Kelly, D. G.; Mackie, A. G.; MacRory, P. P.; McAuliffe, C. A.; Pritchard, R. G.; Watson, S. M. The reaction of coarse-grain metal powders with phosphoranes - A facile activation of metallic reagents and a novel route to metal phosphine complexes - The X-ray crystal structures of MnI₂(PPhMe₂) and FeBr₃(PPhMe₂)₂. *Chem. Commun.* **1991**, 1447–1449.
- (53) Mingos, D. M. P. A General Bonding Model for Linear and Bent Transition Metal-Nitrosyl Complexes. *Inorg. Chem.* **1973**, *12*, 1209–1211. Enemark, J. H.; Feltham, R. D. Principles of Structure, Bonding and Reactivity for Metal Nitrosyl complexes. *Coord. Chem. Rev.* **1974**, *13*, 339–406. Atkinson, F. L.; Blackwell, H. E.; Brown, N. C.; Connelly, N. G.; Crossley, J. G.; Orpen, A. G.; Rieger, A. L.; Rieger, P. H. Synthesis of the 17-electron cations [FeL(L')(NO)₂]⁺ (L, L' = PPh₃, OPPh₃): structure and bonding in four-coordinate metal dinitrosyls, and implications for the identity of paramagnetic iron dinitrosyl complex catalysts. *J. Chem. Soc., Dalton Trans.* **1996**, 3491–3502. Ogasawara, M.; Huang, D.; Streib, W. E.; Huffman, J. C.; Gallego-Planas, N.; Maseras, F.; Eisenstein, O.; Caulton, K. G. RuX(CO)(NO)L₂ and Ru(CO)(NO)L₂⁺: Ru(0) or Ru(II) or In Between? *J. Am. Chem. Soc.* **1997**, *119*, 8642–8651. Jacobsen, H.; Heinze, K.; Llamazares, A.; Schmalle, H. W.; Artus, G.; Berke, H. Coordination chemistry of the [Re(NO)₂(PR₃)₂]⁺ fragment: crystallographic and computational studies. *J. Chem. Soc., Dalton Trans.* **1999**, 1717–1727. Wondimagegn, T.; Ghosh, A. A Quantum Chemical Survey of Metallophosphoryrin-Nitrosyl Linkage Isomers. *J. Am. Chem. Soc.* **2001**, *123*, 5680–5683. McCleverty, J. A. Chemistry of Nitric Oxide Relevant to Biology. *Chem. Rev.* **2004**, *104*, 403–418.
- (54) Rosa, P.; Mézailles, N.; Ricard, L.; Mathey, F.; Le Floch, P.; Jean, Y. Dianionic Iron and Ruthenium(2-) Biphosphinine Complexes: A Formal d¹⁰ Ruthenium Complex with a Square Planar Geometry. *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 1251–1253.
- (55) Perron, H.; Moores, A.; Demachy, I.; Lledos, A.; Jean, Y.; Le Floch, P. Structural flexibility of formally d¹⁰ [M(biphosphinine)₂]^q complexes. *New J. Chem.* **2004**, *28*, 838–842, and references cited.
- (56) Noodleman, L.; Davidson, E. R. Ligand spin polarization and antiferromagnetic coupling in transition metal dimers. *Chem. Phys.* **1986**, *109*, 131–143. Neese, F. Definition of corresponding orbitals and the diradical character in broken symmetry DFT calculations on spin coupled systems. *J. Phys. Chem. Solids* **2004**, *65*, 781–785.
- (57) Louw, W. J.; de Waal, D. J. A.; Kruger, G. J. Preparation of the Five-co-ordinated Complexes [PdX₂(PMe₂Ph)₃] (X = Cl, Br, or I) and the Crystal and Molecular Structure of Dichlorotris-[dimethylphenylphosphine]palladium. *J. Chem. Soc., Dalton Trans.* **1976**, 2364–2368.
- (58) Bartsch, R.; Hietkamp, S.; Morton, S.; Peters, H.; Stelzer, O. Reactions of coordinated ligands. 12. Single-stage template syntheses of tetradentate macrocyclic phosphine complexes. *Inorg. Chem.* **1983**, *22*, 3624–3632.
- (59) Brauer, D. J.; Gol, F.; Hietkamp, S.; Peters, H.; Sommer, H.; Stelzer, O.; Sheldrick, W. S. Reactions of coordinated Ligands. 14. Synthesis of a tetradentate phosphorus macrocycle in a palladium(II) template. *Chem. Ber.* **1986**, *119*, 349–365.
- (60) Addison, A. W.; Rao, T. N.; Reedijk, J.; van Rijn, J.; Vershoor, G. C. Synthesis, Structure, and Spectroscopic Properties of Copper(II) Compounds containing Nitrogen-Sulphur Donor Ligands; the Crystal and Molecular Structure of Aqua[1,7-bis(N-methylbenzimidazol-2'-yl)-2,6-dithiaheptane]copper(II) Perchlorate. *J. Chem. Soc., Dalton Trans.* **1984**, 1349–1356.
- (61) Aullón, G.; Alvarez, S. Axial Bonding Capabilities of Square Planar d⁸-ML₄ Complexes. Theoretical Study and Structural Correlations. *Inorg. Chem.* **1996**, *35*, 3137–3144.
- (62) Deeth, R. J.; Paget, V. J. Molecular mechanics calculations on imine and mixed-ligand systems of Co^{III}, Ni^{II} and Cu^{II}. *J. Chem. Soc., Dalton Trans.* **1997**, 537–546.
- (63) Bernal, I.; Myrczek, J.; Luger, B. J.; Nguyen, M. L. The phenomenon of conglomerate crystallization. XXVI. The crystal and molecular structure of racemic NH₄[Cr(EN)₂OX]Cl₂·H₂O and attempts to prepare NH₄[Co(EN)₂OX]Cl₂·H₂O. *J. Coord. Chem.* **1993**, *29*, 7–16.
- (64) Hua, X.; Larsson, K.; Neal, T. J.; Wyllie, G. R. A.; Shang, M.; Lappin, A. G. Structure and magnetic properties of [Cr(en)₂(ox)][Cr(en)(ox)₂].2H₂O, Δ-[Cr(en)₃]Δ-[Cr(ox)₃] and Δ-[Co(en)(3)]Δ-[Cr(ox)₃]. *Inorg. Chem. Commun.* **2001**, *4*, 635–639.
- (65) Yeo, J. S. L.; Vittal, J. J.; Hor, T. S. A., [PdCl₂(dppfO₂-O,O')]: a simple palladium(II) complex with a rare tetrahedral structure. *Chem. Commun.* **1999**, 1477–1478.
- (66) Orpen, A. G.; Hor, T. S. A.; Vittal, J. J. Personal communication.
- (67) Yamamoto, Y.; Satoh, R.; Tanase, T. Preparation of [Ru(η⁶-C₆Me₆)-Cl(C≡CPh)(RNC)] (R=C₈H₉ or C₆H₂Me₃-2,4,6) and its Reaction with Tetracyanoethylene. Crystal Structures of [Ru(η⁶-C₆Me₆)Cl]{C≡C(CN)₂}-CPh=C(CN)₂}[C₈H₉NC]] and *cis*-[RuCl₂(C₈H₉NC)₂]. *J. Chem. Soc., Dalton Trans.* **1995**, 307–311.
- (68) Drouin, M.; Perreault, D.; Harvey, P. D.; Michel, A. Quasi-One-Dimensional Structure of *cis*-Dichlorobis(2,6-dimethylphenyl isocyanide)-palladium(II), [Pd{2,6-(CH₃)₂C₆H₃NC}Cl₂]. *Acta Crystallogr.* **1991**, *C47*, 752–754.
- (69) Watson, L. A.; Ozerov, O. V.; Pink, M.; Caulton, K. G. Four-Coordinate, Planar Ru(II). A Triplet State as a Response to a 14-Valence Electron Configuration. *J. Am. Chem. Soc.* **2003**, *125*, 8426–8427.
- (70) Martin, R. ePrints UK: Developing a national e-prints archive. <http://www.ariadne.ac.uk/issue35/martin> (22/02/2005). Day, M. EBank UK project scenarios and user requirements. <http://www.ukoln.ac.uk/projects/ebank-uk/requirements/scenarios.html> (22/02/05).
- (71) Schrödinger, L. L. C. *Jaguar, 5.0*; Portland, OR, 2002.
- (72) Schrödinger Inc., *Jaguar, 4.0*; Portland, Oregon, 2000.
- (73) Becke, A. D. Density-Functional Thermochemistry. 3. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. Slater, J. C. *Quantum Theory of Molecules and Solids, Vol. 4: The Self-Consistent Field for Molecules and Solids*; McGraw-Hill: New York, 1974. Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic-Behaviour. *Phys. Rev. A* **1988**, *38*, 3098–3100. Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin-density calculations — a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211. Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron-density. *Phys. Rev. B* **1988**, *37*, 785–789, implemented as described in Miehlich, B.; Savin, A.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.
- (74) Tanner, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicolls, A.; Ringnalda, M.; Goddard, W. A., III.; Honig, B. Accurate First Principles Calculation of Molecular Charge Distributions and Solvation Energies from Ab Initio Quantum Mechanics and Continuum Dielectric Theory. *J. Am. Chem. Soc.* **1994**, *116*, 11875–11882. Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. New Model for Calculation of Solvation Free Energies: Correction of Self-Consistent Reaction Field Continuum Dielectric Theory for Short-Range Hydrogen-Bonding Effects. *J. Phys. Chem.* **1996**, *100*, 11775–11788.
- (75) Glendenning, E. D.; Badenhop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Weinhold, F. *NBO 5.0*; Madison, 2001.

CI0504768