Database Analysis of Transition Metal Carbonyl Bond Lengths: Insight into the Periodicity of π Back-Bonding, σ Donation, and the Factors Affecting the Electronic Structure of the TM−C≡O Moiety

Rosalie K. Hocking*,§ and Trevor W. Hambley*

Centre for Heavy Metals Research, School of Chemistry, University of Sydney, Sydney, NSW 2006, Australia

Received November 23, 2006

Analysis of the relationships between TM−C and C≡O distances of the more than 20 000 structures reported to the CSD has revealed new experimentally derived insights into the bonding in these systems. The databases of structures were investigated by a combination of DFT and statistical methods. The different abilities of transition metals to both accept and donate electrons are reflected in differences in their data sets. There are significant changes in gradient of the curves representing TM−C versus C≡O scatter plots. One such change in gradient occurs at the bond length corresponding to C≡O in its unbound state. There are also significant differences between the second and third transition series. The analysis provides a structural means of probing the distribution of electrons in a metal−carbonyl fragment and provides important insights into the periodicity of back-bonding, σ donation, and the ability of the carbonyl to stabilize different metal oxidation states.

Introduction

When a ligand interacts with a metal, the amount of ligand to metal donation and the metal to ligand back-donation determine many properties of both the metal center (e.g., reduction or oxidation potential) and the ligand (e.g., lability). Understanding these properties is important in understanding a diverse range of chemical properties.[1−4] It is well understood that ligands that can behave as strong π acceptors have the ability to stabilize low metal oxidation states, but determining quantitatively how metal oxidation state and ligand back-bonding interplay has proved experimentally difficult.[5,6]

Experimental probes of metal−ligand bonding in transition metals are usually based on either the metal center or some property of the ligand; the most unambiguous strategies experimentally probe both. Database analysis techniques can provide significant insight into inorganic electronic structure and of a very different nature from that obtained using other techniques, because they enable us to study global trends and the relationship that different geometric fragments have to each other, rather than focusing on either the metal or the ligand, in a single molecule.[7−11]

Transition metal (TM) carbonyl compounds form a class of metal complexes that is one of the most widely studied in chemistry, and of the crystal structures in the CSD, more than 20 000 are transition metal carbonyls.[12−20] These compounds have diverse applications ranging from catalysts to functional species in biochemistry.[12,21] The stability of a transition metal carbonyl bond is a result of the synergistic effects of σ and π bonding, as illustrated in Figure 1.[22,23] σ and π bonding have opposite effects on carbonyl bond order, σ bonding increasing the bond order and π bonding decreasing it. Thus, changes in bonding can be detected by monitoring the carbonyl bond order, usually through the use of IR spectroscopy.[24,25]

There are now so many crystallographically unique transition metal carbonyl observations that these pieces of information

* Author to whom correspondence should be addressed. Tel: 61-2-9351-2830. Fax: 61-2-9351-3329. E-mail: t.hambley@chem.usyd.edu.au.

§ Current address: CSIRO Land and Water, Glen Osmond, South Australia. E-mail: Rosalie.Hocking@csiro.au.

(1) Serres, R. G.; Grapperhaus, C. A.; Bothe, E.; Eckhard, B.; Weyhermueller, T.; Neese, F.; Wieghardt, K. *J. Am. Chem. Soc.* **2004**, *126*, 5138−5153.

(2) Szilagyi, R. K.; Lim, B. S.; Glaser, T.; Holm, R. H.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **2003**, *125*, 9158−9169.

(3) Lever, A. B. P.; Gorelsky, S. I. *Struct. Bonding* **2004**, *107*, 77−114.

(4) Enemark, J. H.; Feltham, R. D. *Coord. Chem. Rev.* **1974**, *13*, 339−406.

(5) Aliaga-Alcalde, N.; Debeer George, S.; Mienert, B.; Bill, E.; Wieghardt, K.; Neese, F. *Angew. Chem., Int. Ed.* **2005**, *44*, 2908−2912.

(6) Walker, F. A. *Inorg. Chem.* **2003**, *42*, 4526−4544.

(7) Hocking, R. K. Ph.D. Thesis, The University of Sydney: Sydney, 2003.

(8) Hocking, R. K.; Hambley, T. W. *Dalton Trans.* **2005**, *5*, 969−978.

(9) Hocking, R. K.; Hambley, T. W. *Chem. Commun.* **2003**, *13*, 1516−1517.

(10) Davis, T. V.; Zaveer, M. S.; Zimmer, M. *J. Chem. Educ.* **2002**, *79*, 1278−1280.

(11) Hunter, A. D. *Organometallics* **1992**, *11*, 2251−2262.

(12) Spiro, T. G.; Zgierski, M. Z.; Kozlowski, P. M. *Coord. Chem. Rev.* **2001**, *219−221*, 923−936.

(13) Vogel, K. M.; Kozlowski, P. M.; Zgierski, M. Z.; Spiro, T. G. *J. Am. Chem. Soc.* **1999**, *121*, 9915−9921.

(14) Cotton, F. A.; Wilkinson, G.; Murillo, C. A.; Bochmann, M. *Advanced Inorganic Chemistry*, 6th ed.; Wiley-Interscience: New York, 1999; p 637.

(15) Lupinetti, A. J.; Strauss, S. H. *Prog. Inorg. Chem.* **2001**, *49*, 1−112.

(16) Willner, H.; Bodenbinder, M.; Broechler, R.; Hwang, G.; Rettig, S. J.; Trotter, J.; von Ahsen, B.; Westphal, U.; Jonas, V.; Thiel, W.; Aubke, F. *J. Am. Chem. Soc.* **2001**, *123*, 588−602.

(17) Aubke, F.; Wang, C. *Coord. Chem. Rev.* **1994**, *137*, 483−524.

(18) Szilagyi, R. K.; Frenking, G. *Organometallics* **1997**, *16*, 4807−4815.

(19) Frenking, G. *J. Organomet. Chem.* **2001**, *635*, 9−23.

(20) Lupinetti, A. J.; Frenking, G.; Strauss, S. H. *Angew. Chem., Int. Ed.* **1998**, *37*, 2113−2116.

(21) Spiro, T. G.; Wasbotten, I. H. *J. Inorg. Biochem.* **2005**, *99*, 34−44.

(22) Dewar, M. J. S. *Bull. Soc. Chim. Fr.* **1951**, *18*, C71−C79.

(23) Chatt, J.; Duncanson, L. A. *J. Chem. Soc.* **1953**, 2939−2947.

(24) Cotton, F. A.; Wilkinson, G. *Advanced Inorganic Chemistry*, 5th ed.; John Wiley and Sons: New York, 1988, p 1034.

(25) Gruhn, N. E.; Lichtenberger, D. L. *Inorganic Electronic Structure and Spectroscopy*; Wiley-Interscience: 1999; Vol. 2, pp 533−574.

10.1021/om061072n CCC: $37.00
Publication on Web 04/17/2007

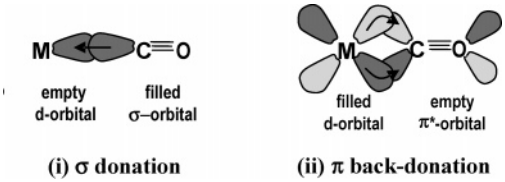

Figure 1. Schematic illustration of the effects of σ donation and π back-donation in transition metal carbonyl compounds.

can be combined in a meaningful way to provide further information about bonding. In this work we examine the nature of the correlation between C≡O and TM−C bond length and its periodicity. The relationships will be discussed in terms of inorganic electronic structure, the energetics of bonding, and the changes in the distribution of electrons throughout the TM−C≡O fragment as its electron population changes.

Experimental Methods

Search Criteria and Data Retrieval. All reported unique terminal TM−carbonyl fragments from crystal structures with *R*-factors < 7.5% were catalogued according to the metal, and the TM−C and C≡O bond lengths were placed in arrays using CSD software.[26,27] Where sufficient data existed, subsets with different coordination numbers, oxidation states, and *trans*-ligands were analyzed.

To identify any trends, three different types of analysis were done. First, principal component analysis (PCA) was performed using the student version of Matlab.[28,29] The resulting line and gradient were then transposed to the average of the data set using Microsoft Excel.[30] As the second means of analysis, the weighted means for the data set were taken; $\bar{x}$(TM−C|C≡O) and $\bar{x}$(C≡O|TM−C). In the first instance the data sets were divided into subsets with different TM−C bond lengths using increments of ∼0.01 Å. For each of these increments the C≡O bond lengths were averaged and the confidence interval (CI) on the mean was calculated as described elsewhere.[9,31,32] In the second instance, the reverse distribution was taken for subsets of different C≡O bond length using increments of ∼0.01 Å. This type of analysis is typically referred to as a locally weighted regression analysis.[33,34]

Excluding Nonterminal Modes of Carbonyl Bonding. In a manner similar to that described previously for carboxylates,[8,31] nonterminal modes of carbonyl bonding were excluded by specifying coordination numbers of the carbon and oxygen atoms. The coordination number of the carbon was specified as two (oxygen and the metal), and that of the oxygen as one (the carbon), thus excluding other carbonyl binding modes, illustrated in Figure S1.

DFT Calculations. Calculations Using Gaussian. A. The starting structures for the three compounds $[Fe(CO)_4]^{2-}$, $[Fe(CO)_5]$, and $[Fe(CO)_6]^{2+}$ were taken from their crystal structures.[35−38] Calculations were performed with the 6-311G* basis set on the Fe atom and 6-31G* on the carbonyl group, as this level of theory/basis set combination has been shown by others to give a well-converged solution.[39−43] The Fe−C bond lengths were set at different values, and the remainder of the molecule was geometry optimized. Thorough studies using similar types of calculations have been reported extensively by Frenking and co-workers.[18,20,44−46] **B.** Starting structures for the compounds $[M(CO)_4]^{2-}$, $[M(CO)_5]$, and $[M(CO)_6]^{2+}$,[47,48] where M = Fe/Ru/Os, were geometry optimized using the basis set SDD and the hybrid functional B3LYP.[49−51] The SDD basis set is standard in the Gaussian98 package and has been used by other authors.[52,53] It treats Ru and Os as 18 valence electron systems with a pseudopotential and employs a [6s,5p,4d/5d] contracted Gaussian for the valence electrons.[54] The remaining atoms in the SDD treatment are assigned a Dunning/Huzinaga[55] valence double-ζ basis set (O[6s,4p] and C[4s,2p]). A double-ζ basis set was thought to suffice for this work, as the interest was in geometry. All geometries were optimized with the default criteria in Gaussian98.

Calculations Performed Using ADF. Gradient-corrected calculations were performed using the exchange functional of Becke[50] and the correlation functional of Perdew[50,56] (BP86). The frozen core approximation[57] was used for the 1s−4d orbitals of Os, the 1s−3d orbitals of Ru, the 1s−2p orbitals of Fe, and the 1s orbitals of O and C. Scalar relativistic corrections that used the zero-order regular approximation (ZORA)[58−60] were applied to all ADF calculations. For valence orbitals, Slater-type orbital (STO) basis sets of triple-ζ quality were employed with polarization functions on the ligand atoms (2p for H, 3d for all others) and additional

(26) Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taymor, R. *Acta Crystallogr.* **2002**, *B58*, 389−397.
(27) *Quest 3D*, a program for searching the CSD; CCDC: 12 Union Road, Cambridge, UK, 1994.
(28) *Matlab*, 1.6 ed.; Math Works Inc., 2002.
(29) The principal component is by definition the first eigenvalue of the covariance matrix, in Matlab: {[*a*,*b*] = eig(cov(2*x*(no. of obs.) matrix))}.
(30) *MicroSoft Excel*; Microsoft Inc., 2000.
(31) Hocking, R. K.; Hambley, T. W. *Inorg. Chem.* **2003**, *42*, 2833−2835.
(32) Hocking, R. K.; Hambley, T. W. *Inorg. Chem.* **2002**, *21*, 2660−2666.
(33) Spiegel, M. R. *Probability and Statistics*; McGraw-Hill Book Company: New York, 1975.
(34) Harris, R. J. *A Primer of Multivariate Statistics*; Academic Press: New York, 1985.
(35) Teller, R. G.; Finke, R. G.; Collman, J. P.; Chin, H. B.; Bau, R. *J. Am. Chem. Soc.* **1977**, *99*, 1104−1111.
(36) Chin, H. B. *J. Am. Chem. Soc.* **1976**, *98*, 2434−2439.
(37) Braga, D.; Grepioni, F.; Orpen, A. G. *Organometallics* **1993**, *12*, 1481−1483.
(38) Bernhardt, E.; Bley, B.; Wartchow, R.; Willner, H.; Bill, E.; Kuhn, P.; Sham, I. H. T.; Bodenbinder, M.; Brochler, R.; Aubke, F. *J. Am. Chem. Soc.* **1999**, *121*, 7188.
(39) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. *J. Chem. Phys.* **1998**, *109*, 1223−1229.
(40) McGrath, M. P.; Radom, L. *J. Chem. Phys.* **1995**, *103*, 6104−6113.
(41) Curtiss, L. A.; McGrath, M. P.; Blaudeau, J-P.; Davis, N. E.; Binning, R. C. J.; Radom, L. *J. Chem. Phys.* **1995**, *103*, 6104−6113.
(42) Ryde, U.; Olsson, M. H. M.; Pierloot, K. In *Theoretical and Computational Chemistry*; Elsevier: 2001; Vol. 9, pp 1−55.
(43) Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421−437.
(44) Frenking, G.; Wichmann, K.; Frohlich, N.; Loschen, C.; Lein, M.; Frunzke, J.; Rayon, V. M. *Coord. Chem. Rev.* **2003**, *238−239*, 55−82.
(45) Ehlers, A. W.; Frenking, G. *J. Am. Chem. Soc.* **1994**, *116*, 1514−1512.
(46) Ehlers, A. W.; Frenking, G. *Chem. Commun.* **1993**, 1709−1710.
(47) Finze, M.; Bernhardt, E.; Willner, H.; Lehmann, C. W.; Aubke, F. *Inorg. Chem.* **2005**, *44*, 4206−4214.
(48) Bernhardt, E.; Back, C.; Bley, B.; Wartchow, R.; Westphal, U.; Sham, I. H. T.; von Ahsen, B.; Wang, C.; Willner, H.; Thompson, R. C.; Aubke, F. *Inorg. Chem.* **2005**, *44*, 4189−4205.
(49) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(50) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(51) Becke, A. D. *J. Chem. Phys.* **1986**, *84*, 4524−4529.
(52) Iron, M. A.; Lo, H. C.; Martin, J. M. L.; Keinan, E. *J. Am. Chem. Soc.* **2002**, *124*, 7041−7054.
(53) Toh, J. S. Honours Thesis; The University of Sydney, 2000.
(54) Andrae, D.; Haussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, *77*, 123−129.
(55) Dunning, T. H. J.; Hay, P. J. In *Modern Theoretical Chemistry*; Schaefer, H. F., Ed.; Springer: New York, 1976; Vol. 3, pp 1−123.
(56) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822−8824.
(57) Baerends, E. J.; Ellis, D. E.; Ros, P. *Theor. Chim. Acta* **1972**, *27*, 339−354.
(58) van Lenthe, E.; Baerends, E. J.; Snijders, J. G. *J. Chem. Phys.* **1993**, *99*, 4597−4610.
(59) van Lenthe, E.; Baerends, E. J.; Snijders, J. G. *J. Chem. Phys.* **1994**, *101*, 9783−9790.
(60) van Lenthe, E.; Ehlers, A. E.; J., B. E. *J. Chem. Phys.* **1999**, *110*, 8943−8950.

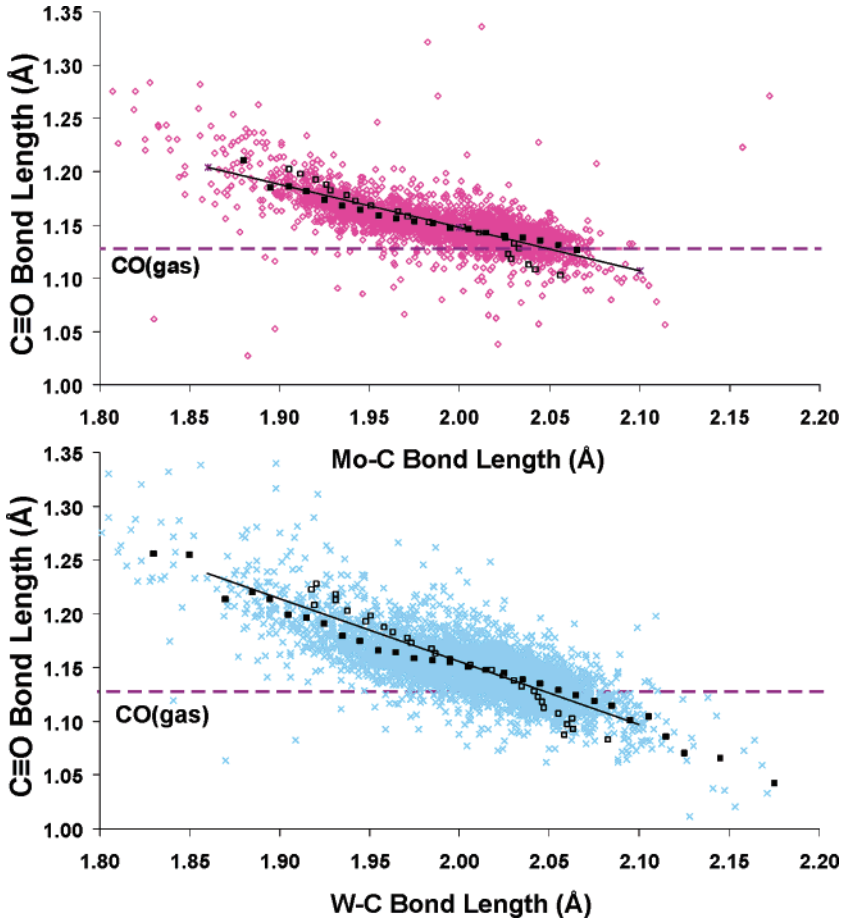

Figure 2. Scatter plot of C≡O bond length vs TM−C bond length for Mo and W. Superposed are the weighted averages, the mean C≡O bond length for a given TM−C bond length {$\bar{x}$(C≡O|TM−C)} (■), $\bar{x}$(TM−C|C≡O) (□), and the principal component analysis (—).

valence p orbitals on the metal atoms, i.e., ADF basis set IV.[61,62] This basis set combination was chosen, as previous studies have shown that it gives a well-converged solution.[42,43]

Results and Analysis

I. Preliminary Analysis of TM−Carbonyls and the Correlation between TM−C and C≡O Bond Lengths. Several different analysis techniques were applied in order to extract useful insights from the relationship between TM−C and C≡O bond lengths. Figure 2 and Figures S3−S5 represent superpositions of the scatter plots, the principal component axis, and the locally weighted averages for the distributions of TM−C versus C≡O bond lengths for each of the elements Mo and W (Figure 2) groups 6 and 7 (Figure S2), groups 8 and 9 (Figure S3), and group 10, V and Cu (Figure S4).

If the scatter plots represented a single relationship between TM−C and C≡O bond lengths, PCA would give the best description of any linear relationship between them. However, these data sets do not represent a single relationship, but the sum of many relationships. Thus, the weighted regression is likely to be the most useful measure, as it provides insight into the data set behavior, and the curves representing the weighted regression $\bar{x}$(C≡O|TM−C) (Figures 2 and 4) gave the most useful information about TM−carbonyl bonding. Also the range of C≡O bond lengths is small, compared to the observed range

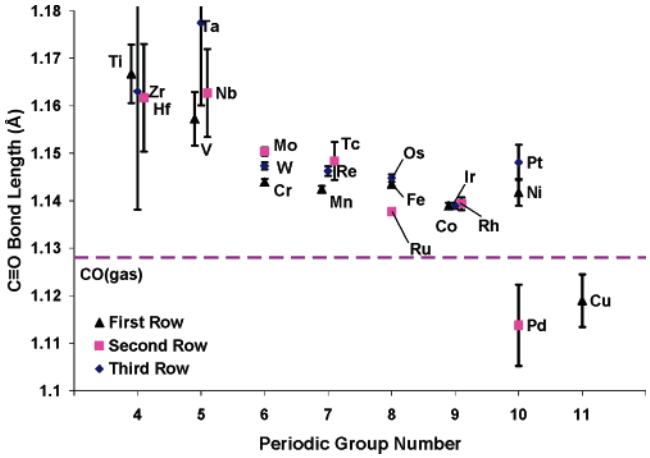

Figure 3. Plot of average C≡O bond length vs periodic group number for all transition metals where TM−carbonyl species have been reported. Error bars represent the 95% confidence intervals on the mean.

of TM−C bond lengths. Indeed for the Mo data set (Figure 2) the C≡O bond length range is not much greater than what we would expect to arise from crystallographic error alone. The average standard deviation on C≡O bond lengths calculated for given Mo−C bond lengths between 1.90 and 2.05 Å is 0.012 Å.[65−67] The same shaped curve is observed for the conditional

(61) Vernooijs, P.; Snijders, G. P.; Baerends, E. J. *Slater Type Basis Functions for the Whole Periodic System*, Internal Report; Free University: Amsterdam, The Netherlands, 1981.
(62) Snijders, J. G.; Vernooijs, P.; Baerends, E. J. *At. Data Nucl. Data Tab.* **1981**, *26*, 483−509.
(63) Hurlburt, P. K.; Rack, J. J.; Luck, J. S.; Dec, S. F.; Webb, J. D.; Anderson, O. P.; Strauss, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 10003−10014.
(64) Dias, H. V. R.; Jin, W. *Inorg. Chem.* **1996**, *35*, 3687−3688.

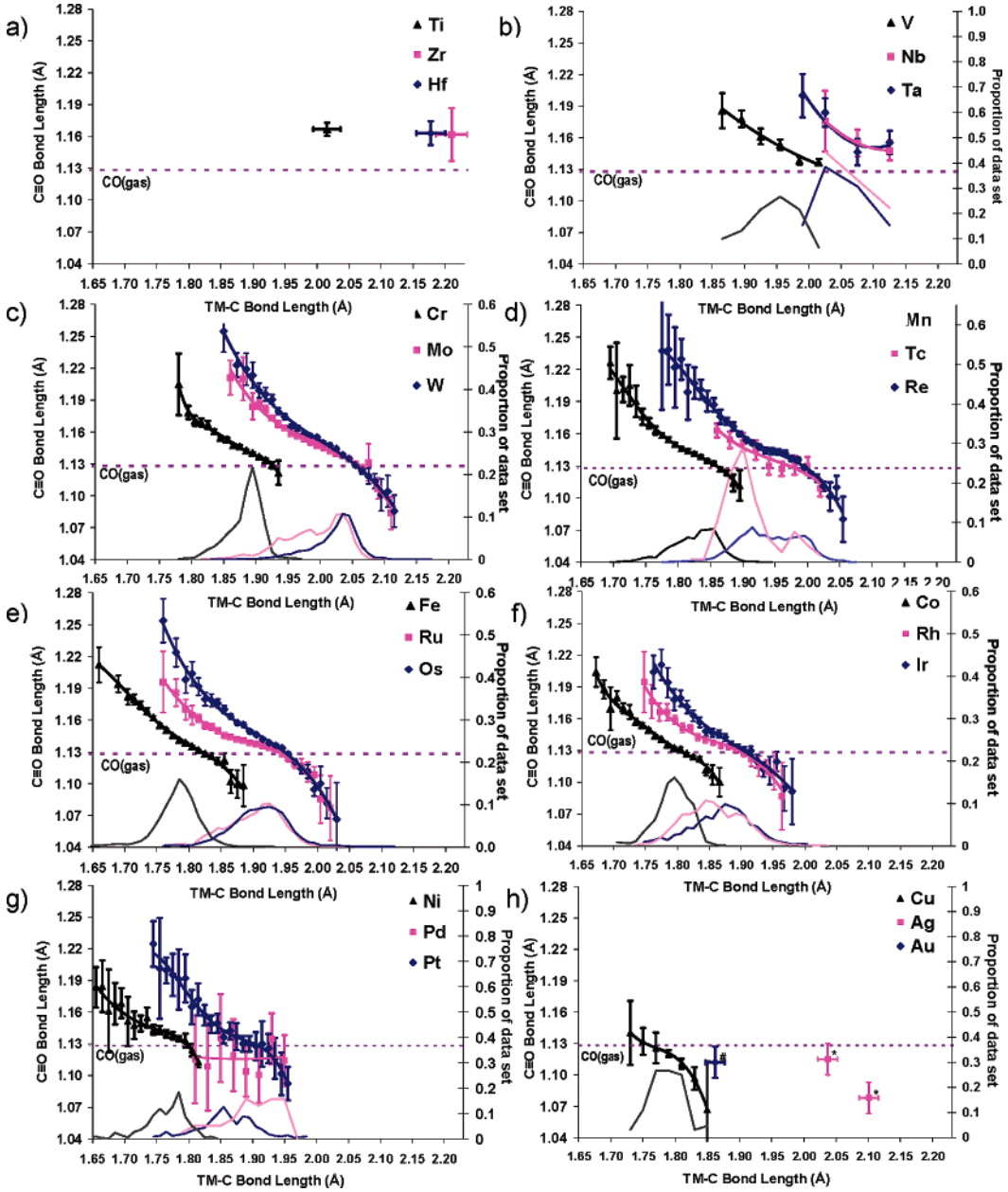

Figure 4. Plot of the conditional $\bar{x}$(C≡O|TM−C) bond length (Å) for groups 4−11: (a) group 4 (6c data), (b) group 5 (6c data), (c) group 6 (6c data), (d) group 7 (6c data), (e) group 8 (6c data), (f) group 9 (all data), (g) group 10 (all data), and (h) group 11 Cu, (all data)Ag*,[63] and Au #.[64] The curves under the data sets represent the proportion of the data set at each point.

averages in each direction; that is, the $\bar{x}$(C≡O|TM−C) curve has the same chair shape as the $\bar{x}$(TM−C|C≡O) curve, even though the likely error on the latter is significantly greater (Figures S2−S4), since the scatter is greater. These observations justify the use of the weighted averages as the primary analysis tool and the analysis of structures as a group even though each individual observation is subject to a variety of influences such as *trans* influences, changes in oxidation state, and geometry. Further, these effects can be analyzed by taking subsets of structures, and this makes only a small difference to the geometric trends shown in the full data sets (*trans* influences (Figure S14), different oxidation states (Figures S15, S16), the effect of coordination number (Figure S17)).

Several "expected" observations emerge from the plots in Figures 2−4. First, shorter TM−C bonds correspond to longer C≡O bonds. This is expected on the basis of both experimental[12,13,16,17,68−72] and theoretical work[18,20,73−75] and is con-

(65) Orpen, A. G.; Quayle, M. J. *J. Chem. Soc., Dalton Trans.* **2001**, 1601−1610.
(66) Orpen, A. G. *Acta Crystallogr.* **2002**, *B58*, 398−406.
(67) Orpen, A. G. *Chem. Soc. Rev.* **1993**, 191−197.
(68) Kettle, S. F.; Diana, E.; Stanghellini, P. L. *Inorg. Chem.* **1998**, *37*, 6502−6510.
(69) Nemscsok, D. S.; Kovacs, A.; Rayon, V. M.; Frenking, G. *Organometallics* **2002**, *21*, 5803−5809.
(70) Kettle, S. F.; Diana, E.; Boccaleri, E.; Stanghellini, P. L. *Eur. J. Inorg. Chem.* **1999**, 1957−1963.
(71) Bernhardt, E.; Bley, B.; Wartchow, R.; Willner, H.; Bill, E.; Kuhn, P.; Sham, I. H.; Bodenbinder, M.; Brochler, R.; Aubke, F. *J. Am. Chem. Soc.* **1999**, *121*, 7188−7200.
(72) Brunet, J.-J.; Chauvin, R.; Diallo, O.; Kindela, F.; Leglaye, P.; Neibecker, D. *Coord. Chem. Rev.* **1998**, *178−180*, 331−351.
(73) Sherwood, D. E.; Hall, M. B. *Inorg. Chem.* **1983**, *22*, 93−100.
(74) Bridgeman, A. J. *Inorg. Chim. Acta* **2001**, *321*, 27−21.
(75) Ziegler, T.; Tschinke, V.; Ursenbach, C. *J. Am. Chem. Soc.* **1987**, *109*, 4825−4837.

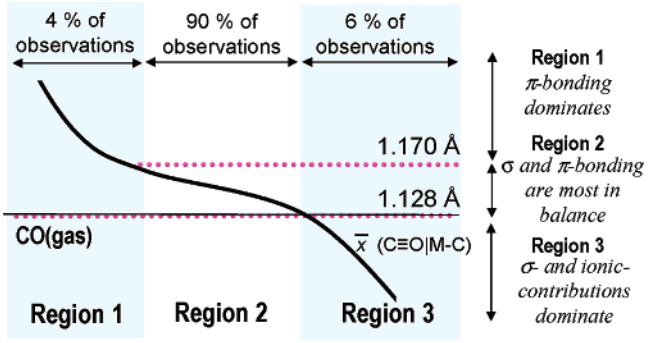

Figure 5. Schematic illustration based on Figure 4, showing the regions where different types of bonding dominate.

sistent with the expected effects of increasing π overlap as the TM−C bond length decreases. Second, the bond length trends are periodic; TM−C bond lengths decrease across and increase down the periods as expected on the basis of the covalent radii of the transition metals[76] (Figures 3 and 4, S5−S7). Third, we observe a general decrease in C≡O bond length and hence π bonding upon going from left to right across the periodic table (Figure 3). This is expected because the d orbital energy decreases on average going across the period, making π bonding less favorable.[77,78]

Two other features are evident in Figures 2 and 4: first, all the curves have an inflection and, second, there are differences in the behavior of the curves for second- and third-row transition metals, the second-row curve exhibiting a larger inflection and smaller gradient than the third. Previously,[9] we suggested that these data sets span three regions: the longest C≡O bonds and shortest TM−C bonds occupy region 1, where π bonding dominates over σ, intermediate TM−C and C≡O bond lengths occupy region 2, where σ and π bonding are more in balance, and the shortest C≡O bond lengths occupy region 3, where σ and ionic contributions to bonding dominate. A schematic showing these regions is given in Figure 5. The curves for the second- and third-row transition elements (Figure 4) overlay one another in region 3, but increasingly diverge though regions 1 and 2 for all triads except Cu.

II. Analysis of Homoleptic TM−Carbonyl Species and Mutual Ligand Effects. Attaching meaning to the correlations found in crystallographic data sets can be complicated because any given data set contains information from many different compounds. Thus, when comparing a group of observations, the dominant variations across the data set must be identified. In the present study we must identify the dominant determinant of TM−C bond lengths. Any particular TM−ligand bond length is the consequence of the ionic and covalent contributions to bonding. However, in the TM−carbonyl data sets, several features of the data sets indicate that covalent contributions to bonding are the dominant determinants of TM−C distance. First, each scatter plot has a continuous, well-defined shape, determined by the fact that as the TM−C bond gets longer, the C≡O bond gets shorter. This continuous shape is indicated by the black outline around the cobalt carbonyl data set shown in Figure 6. Many data sets of transition metal-ligand bond lengths plotted this way do not exhibit a continuous and well-defined shape.[79] For comparison, scatter plots of Co−O distances versus

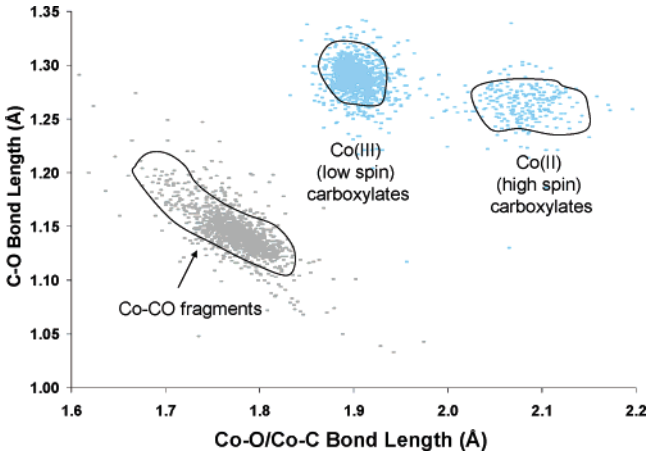

Figure 6. Comparison of the data set structure of cobalt carbonyls and cobalt carboxylates.[8,31]

O−C distances of the monodentate carboxylates[8,31] reveal two clusters of data: the cluster with the shortest Co−O distances represents low-spin Co(III) and the cluster with the shortest Co−O distances represents high-spin Co(II). We have previously[8,31] reported that the covalency of these systems is correlated with the O−C carboxylate distance, and therefore, if metal−ligand bond length was a meaningful indicator of metal−ligand covalency, we would expect the data sets to exhibit a correlation between the Co−O and O−C bond lengths. The fact that these parameters are not significantly correlated indicates that we cannot state definitively what the major determinate of Co−O bond length is.[80] In the case of the carbonyls, if ionic bonding had a larger effect on TM−C bond length, we would expect a different type of scatter plot: one in which the TM−C and O−C bond lengths were not so highly correlated. Thus, the primary determinant of the TM−C bond lengths is likely to be covalent contributions to bonding, whether they arise from σ donation or π back-donation.

To further establish the determinants of the TM−C bond lengths, subsets of the larger data sets were considered and DFT calculations were undertaken. Homoleptic TM−carbonyl compounds are an important class of transition metal compounds because they are not altered by mutual ligand effects. This enables an analysis of the effect of a change in oxidation state on the TM−C versus C≡O relationship without the complication of contributions from mutual ligand effects. From the data for the homoleptic transition metal carbonyl complexes given in Figure 7, it is clear that in this group the primary determinant of C≡O bond length is oxidation state. The same conclusion has previously been derived from an analysis of IR frequencies.[81]

Figure 8 shows the average bond lengths for homoleptic TM−carbonyl complexes superposed on the $\bar{x}$(C≡O|TM−C) curves for the first-row transition metals and reveals that the average bond length for each of the homoleptic transition metal carbonyl groups lies within error of the respective curves. An analysis of the homoleptic species alone could lead to the conclusion that the variation that occurs across the data sets represents a change in oxidation state, since in all cases the C≡O bond length is to a first approximation determined by oxidation state. However, the majority (Table S5) of observations of TM−carbonyl bonds represent complexes in oxidation state 0 or 1, and the curves in their entirety can be reproduced from these complexes alone (Figures S18−S20). This implies that a change

(76) Shannon, R. D. *Acta Crystallogr.* **1976**, *A32*, 751−767.
(77) Pauling, L. *The Nature of the Chemical Bond*; Cornell University Press, 1960.
(78) Pearson, R. G. *Acc. Chem. Res.* **1990**, *23*, 1−2.
(79) Another example of this is given in Figure S19, where the Fe−CO fragments are compared to the Fe−NO fragments.
(80) It is worth noting that many other transition metal carboxylate data sets do exhibit a negative correlation.
(81) Willner, H. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 2402−2425.

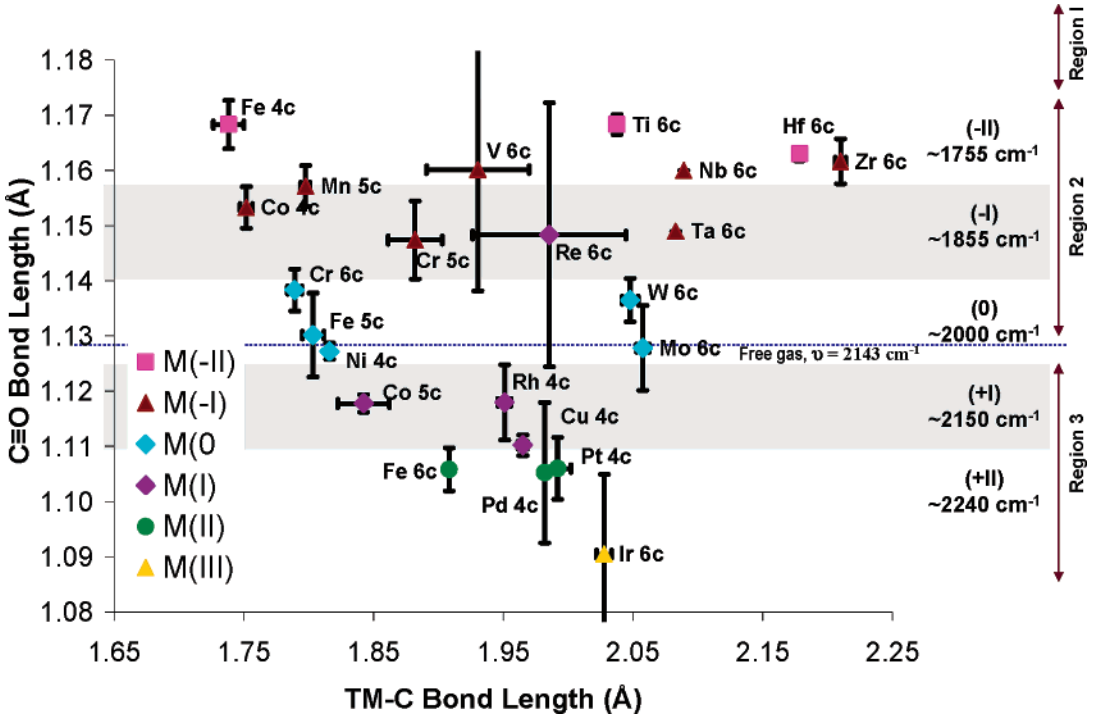

Figure 7. Structural data reported for homoleptic TM−carbonyl species illustrating the relationships between C≡O bond length, C≡O stretching frequency, and metal oxidation state.[81]

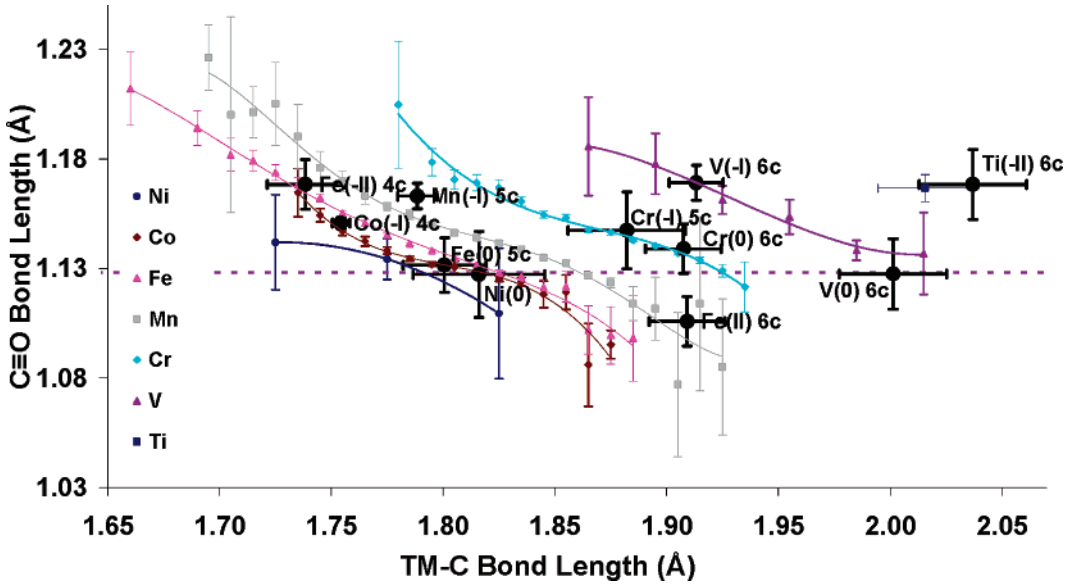

Figure 8. Average bond lengths for homoleptic TM−carbonyl complexes superposed on the C≡O vs TM−C bond length trends for the first-row transition metals.

of ligands in the coordination sphere can have a similar effect on the C≡O bond length to a change in oxidation state. This observation is fundamental to the interpretation of these relationships because it reveals that the variation across the data sets is due to a change in the electron population of the TM−C≡O fragment, whether this arises from a change in formal oxidation state or from mutual ligand effects. The results indicate that those fragments with the shortest TM−C bond lengths and longest C≡O bond lengths have the highest electron population (most reduced), and those fragments with the longest TM−C bond length and shortest C≡O bond length have the lowest electron population (most oxidized).

To further complement the analysis of the subsets of crystallographic data, DFT calculations were performed. First, the Fe−C versus C≡O relationships for the complexes $[Fe(CO)_4]^{2-}$, $[Fe(CO)_5]$, and $[Fe(CO)_6]^{2+}$ were examined using DFT calculations in which the Fe−C bond length was fixed and the C≡O bond length was geometry optimized (Figure 9). Similar calculations have been reported previously.[18,45,46,83] When the Fe−C bond lengths in each of the three compounds were systematically varied, only small changes in C≡O bond length relative to the changes across the data sets were observed (Figure 9). This indicates that the distribution of the bond lengths in the crystallographically determined structures relates not to any inherent relationship between TM−C and C≡O distance for a given TM in a particular oxidation state, but rather to differences in the effective oxidation state.

(82) Herzberg, G. *Molecular Spectra and Molecular Structure*; Van Nostrand Reinhold: New York, 1966.
(83) Frenking, G.; Frohlich, N. *Chem. Rev.* **2000**, *100*, 717−774.

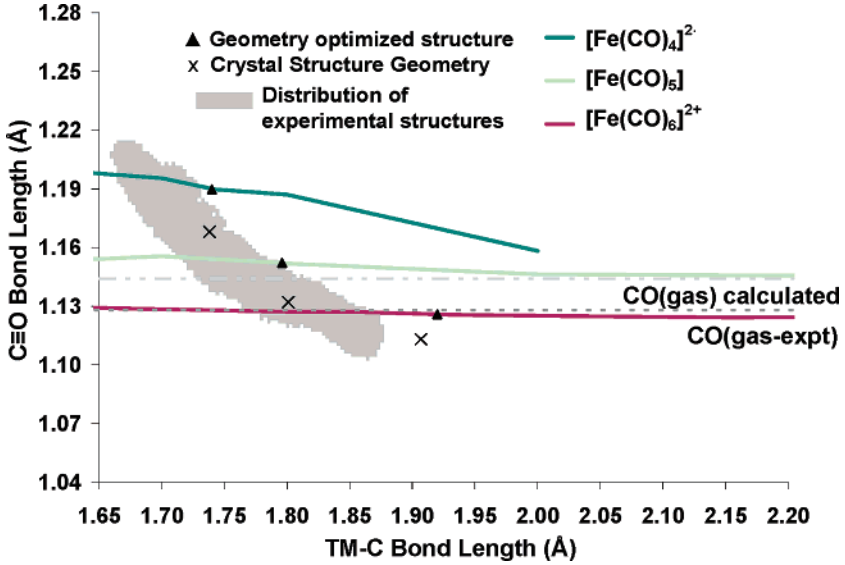

Figure 9. Fe−C vs C≡O relationships derived by taking the molecules $[Fe(CO)_4]^{2-}$, $[Fe(CO)_5]$, and $[Fe(CO)_6]^{2+}$ through a symmetric stretch of the Fe−C bonds. Superposed are the experimental structures (×), optimized geometry (▲) (lowest energy point on the Fe−C vs C≡O curve), and the distribution of Fe−C≡O observations (shaded in gray).

Table 1. Calculated and Observed Bond Length Parameters for $[M(CO)_4]^{2-}$, $[MCO_5]^0$, and $[MCO_6]^{2+}$, where M = Fe, Ru, and Os[a]

	source	Fe−C	C≡O	Ru−C	C≡O	Os−C	C≡O
$[M(CO)_4]^{2-}$	B3LYP	1.744	1.227	1.893	1.225	1.907	1.228
	BP86	1.748	1.199	1.910	1.194	1.909	1.208
	expt	1.738	1.168	N/A	N/A	N/A	N/A
$[M(CO)_5]^0$	B3LYP	1.807	1.175	1.950	1.177	1.967	1.166
	BP86	1.887	1.132	1.954	1.155	1.960	1.158
	expt	1.801	1.132	N/A	N/A	N/A	N/A
$[M(CO)_6]^{2+}$	B3LYP	1.931	1.148	2.029	1.149	2.044	1.150
	BP86	1.894	1.132	2.033	1.132	2.035	1.132
	expt	1.907[71]	1.113	2.023[47,48]	1.101	2.010[47,48]	1.103

[a] Values were obtained using DFT (B3LYP and BP86) or from the crystal structure (expt). Free C≡O bond length (B3LYP/SDD = 1.167 Å, BP86/ADF/BSIV (ZORA) = 1.140 Å, expt 1.128 Å).[82]

A second set of calculations was performed to examine the effects of periodicity. The geometries of nine compounds ($[M(CO)_4]^{2-}$, $[M(CO)_5]^0$, and $[M(CO)_6]^{2+}$, where M = Fe, Ru, or Os) were calculated by DFT, and the results of these calculations are given in Table 1. These complexes were chosen because a complete set of crystallographic observations exist for M = Fe and for $[M(CO)_6]^{2+}$ for M = Fe, Ru, or Os. Additionally, it was expected that the large range of oxidation states would encompass a range of TM−ligand bond lengths and bonding arrangements.

Of the geometry-optimized structures in Table 1, all of the Fe structures and the $[Ru(CO)_6]^{2+}$ and $[OsCO)_6]^{2+}$ structures are in accord with the experimental data and fit reasonably well with the trend lines.[35−38,47,48] The geometry-optimized structures of the remaining four compounds do not fit close to the TM−C versus C≡O relationships in Figure 4. Further, the Ru and Os compounds of the same oxidation state have similar geometries, i.e., metal−ligand bond lengths and C≡O bond lengths. In contrast the scatter plots for highly reduced Ru and Os species diverge in regions 1 and 2. Since DFT calculations for almost all transition metal carbonyls that have been reported compare well with the experiment,[20,45,74,84] and different functionals give similar results, it is likely that DFT gives a good estimation of the geometric structure of the three compounds for which no experimental data exist. Data sets of crystallographic data do not represent a random sample of all possible compounds, but represent those compounds that are stable enough to have been crystallographically characterized. Thus, there are significant and quantifiable biases between the structures for the second transition series and the third transition series. The fact that the optimized structure of $[Ru(CO)_4]^{2-}$ does not fit on the general trend lines suggests that it is not typical of Ru−carbonyl compounds that have been reported to the CSD.

Discussion

The shape of the curves here arises from a combination of chemical and statistical factors. To understand how statistical factors such as the proportion of structures at each point (frequency distribution) affect the conditional expectation value curves, we will first consider a hypothetical data set constructed so that the underlying relationship between two variables A and B is $A = B$. The frequency distribution of the data set is centered about the observation $A = \alpha$ and $B = \beta$ and is given by the pale purple curve in Figure 10. In the limit where there is no error on an observation $\{A,B\}$ the conditional expectation value curve $\bar{x}(A|B)$ will produce the relationship $A = B$. However, where there is an error component, the conditional expectation values become "contaminated" by the more frequent observations such that both the points $A = \alpha + \delta$ and $A = \alpha - \delta$ would have a contamination of points from $A = \alpha$, as shown in Figure 10. This systematic contamination of observations would alter the $\bar{x}(A|B)$ distribution so as to give an inflection about the most frequent observation, shown by the black curve in

(84) Gray, H. B.; Beach, N. A. *Inorg. Chem.* **1963**, *85*, 2922−2927.

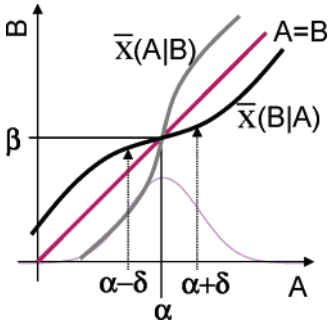

Figure 10. Statistical observation of a relationship $A = B$ in which the frequency of observations is centered about the point $A = \alpha$ and $B = \beta$. The relationship $A = B$ is given by the purple line, the conditional expectation value $\bar{x}(A|B)$ curve is given in black, and $\bar{x}(B|A)$ is given in gray.

Figure 10. If the average error on the observation of variables A and B were the same, then we would expect the calculation of the $\bar{x}(B|A)$ curve to give the reverse inflection, as shown by the gray line in Figure 10, and the principle component axis to return the relationship $A = B$.

The data sets of TM−C versus C≡O bond lengths considered herein are more complicated than the case shown in Figure 10 for two reasons. First, how often a fragment is observed is related to its chemical properties; and second, a database of fragments does not necessarily have any underlying relationship such as $A = B$. It was shown in the analysis section that PCA gives an unsatisfactory description of the data sets described here. It is clear from an examination of Figure 2 that the principle component fails to reflect parts of the W and Mo data sets, implying that the error associated with an observation {TM−C, C≡O} is smaller than the underlying chemical factors, which may bias the data sets and produce an effect on $\bar{x}$(C≡O|TM−C).

To investigate the relationship between how often something is observed and the conditional expectation values, the proportion of the data set as a function of TM−C distance was calculated and is given by the lower curves in Figure 4. The right axis on each graph gives the proportion of the data set such that the area under each curve integrates to 1. This allows a direct comparison of the data sets for each transition metal. The proportion of the data set as a function of C≡O bond length is given in Figures S5 and S6.

It is clear that the proportion of the data sets at each point is different depending on the transition metal. However, for all data sets the same general trend is reproduced, when the observations are dominated by long C≡O distances (V triad) or short ones (Cu and Ni triads). Thus the differences in the $\bar{x}$(C≡O|TM−C) relationship must reflect differences in bonding that systematically bias the observations.

This observation is consistent with the other aspects of the curves that have chemical implications; for example, the curves of the second and third transition series are seen to increasingly diverge for C≡O distances longer than the free gas bond length. We can select structures and accurately reproduce the structure with DFT (*vide supra*).

Thus, we need to consider factors other than error that may cause a systematic bias in the data sets. If structures are biased to lower energy, then the bonding differences between the second and third transition series will account for their geometric differences. To consider this further, the effect of changing TM on CO bonding is considered by MO theory.

Calculations have shown that the 5d orbital is destabilized by the relativistic contraction of the inner orbitals, and consequently the 5d orbital is effectively higher in energy than the 4d.[85,86] The likely effect of this is best visualized using an MO picture in O_h symmetry shown in Figure 11. When a metal interacts with a ligand, the strength of the overlap is determined by how far apart in energy the orbitals are (ΔE) and the overlap integral S. Since any energetic factors that stabilize or destabilize carbonyl orbitals are the same for all data sets and S is approximately proportional to bond length, it is only effects relating to the relative energy of the transition metal orbitals need be considered.[84] For an equivalent interaction the difference in energy between the 5d orbitals and the π acceptor orbitals on the carbonyl is less than that between the 4d and the same set of orbitals (i.e., ΔE_π 5d < 4d, Figure 11) favoring stronger π back-donation. The reverse is true when σ donation is considered (ΔE_σ 4d < 5d) because the lowering of the 4d orbitals decreases the energy difference between the carbonyl σ donor orbitals, making σ overlap more favorable, explaining the divergence one observes between the second and third transition series data sets, once π bonding starts to have an effect on the C≡O bond length.

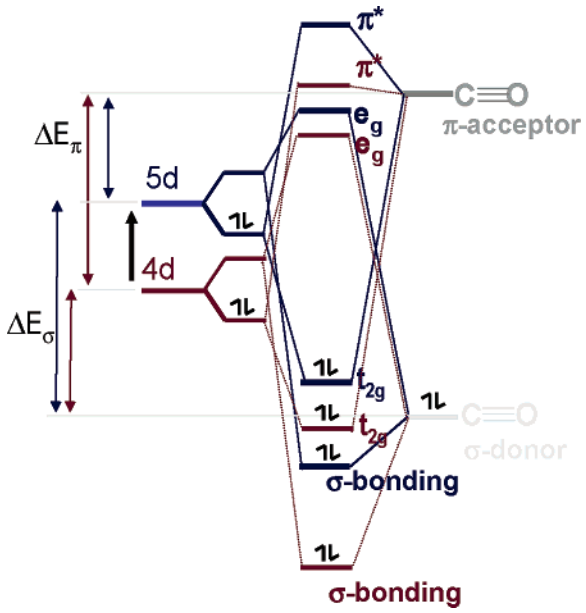

Figure 11. Molecular orbital diagram representing the interaction of carbonyl σ donor and π acceptor orbitals with both 4d and 5d transition metals. Upon going from 4d to 5d ΔE_π gets smaller and ΔE_σ gets larger, decreasing π back-bonding and increasing σ donation.

The origin of the divergence of the plots for the second and third transition periods also allows us to more generally understand the chair shape of all the data sets. If we examine bonding across the data sets as electrons are added to the TM−C≡O fragment, they do not add equally to the metal and to the carbonyl. In the middle of the data sets the lower gradient indicates that there is a tendency for the electrons to add more to the metal than to the carbonyl, and at each end to add more equally to both the metal- and carbonyl-based orbitals. This tendency could be considered to be a reflection of the energetics of electron distribution within the TM−C≡O fragment, the curves reflecting the energetically most favorable distribution of electrons in a TM−C≡O fragment.

(85) Li, J.; Schreckenbach, G.; Ziegler, T. *Inorg. Chem.* **1995**, *34*, 3245−3252.
(86) Li, J.; Schreckenbach, G.; Ziegler, T. *J. Am. Chem. Soc.* **1994**, *117*, 486−494.

Summary

In this study we have utilized a large number of structures to gain new insights into both the periodicity of transition metal carbonyl bonding and the energetic factors determining metal- versus ligand-based oxidation and reduction. While we cannot separate the effects of different oxidation states, we can examine how the fragments in the CSD demonstrate a preference for particular TM−C≡O geometric combinations from which we can infer a distribution of the electrons. The trends derived allow us to quantitate the effects of different metals on bonding phenomena and to more generally understand the periodicity of σ donation, π back-donation, and the factors affecting the stability of different metal and ligand oxidation states.

Acknowledgment. We would like to thank the Australian Centre for Advanced Computing and Communications (ac3) for access to computer time. Professors Peter A. Lay and Hans B. Bürgi and Dr. Jim Eckert are thanked for useful discussions. R.K.H would like to thank the Australian government for an APA, and Dr. Joan R. Clark for a travel scholarship.

Supporting Information Available: Scatter plots equivalent to Figure 2 for all transition metal data sets, data sets plotted with the opposite conditional expectation values $\{\bar{x}(TM-C|C\equiv O)\}$, and the relative proportions of C≡O bond lengths; scatter plots showing standard deviations for all data sets, $\{\bar{x}(C\equiv O|TM-C)\}$ values plotted for the first, second, and third row; a more detailed analysis of subsets including the effects of *trans*-ligands, metal−metal bonds, and coordination number; and a more detailed discussion of error analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

OM061072N