ARTICLES

Modified Carbon Pseudopotential for Use in ONIOM Calculations of Alkyl-Substituted Metallocenes†,‡

John L. Lewin and Christopher J. Cramer*

Department of Chemistry and Supercomputer Institute, University of Minnesota, 207 Pleasant Street SE, Minneapolis, Minnesota 55455

Received: December 17, 2007; Revised Manuscript Received: March 04, 2008

Nonrelativistic carbon 1s core pseudopotentials are optimized for substituted cyclopentadienide ring carbons for use in integrated molecular orbital molecular orbital (IMOMO) and integrated molecular orbital molecular mechanics (IMOMM) calculations where the Cp ring substituents are not included in the high-level IMOMO or IMOMM subsystem. Use of the optimized pseudopotential within the IMOMO framework leads to significant improvements in predicted carbonyl stretching frequencies for a series of Cp-ring-methylated zirconocenes compared to using a standard carbon pseudopotential. The technology is less successful in the IMOMM implementation.

Introduction

Metallocenes are a well characterized class of compounds involving a metal supported by two cyclopentadienide ($C_5H_5^-$, Cp) ligands in what is commonly called a sandwich complex.[1] The term metallocene has also come to include informally arene−metal complexes with ligands such as benzene, naphthalene, and cyclooctatetraene, in numbers ranging from the traditional two (sandwich), to only one (half-sandwich), to as many as three or more. Select metallocenes are particularly noteworthy for their roles as polymerization catalysts and selective C−H bond activators in organic synthesis.[1–5] Some of the strongest C−H bond activators are the group III metallocenes, including scandium, yttrium, and lutetium.[4–7] As the active species do not require ill-defined alkyl-alumoxane cocatalysts, nor do they suffer from the ion-pairing complications associated with cationic alternatives (e.g., zirconocene or titanocene catalysts), group III metallocenes are particularly suited to mechanistic studies of key polymerization steps, such as chain initiation, propagation, and transfer. The tendency for group III metallocenes to dimerize when the supporting ligand is Cp has led to the widespread use of pentamethylcyclopentadienide ($C_{10}H_{15}^-$, Cp*), whose increased steric bulk facilitates the isolation of well-behaved, monomeric, single-metal complexes.[3,4,8] These Cp^*_2M species generally have the added benefit of increased stability and solubility.

In addition to experimental studies of group III metallocenes, a substantial number of computational modeling efforts have been undertaken to understand various aspects of the structure and reactivity of these species.[9–34] The application of quantum mechanical methods to study the relevant chemistry involves the selection of various modeling protocols. In particular, while an unsubstituted metallocene having two Cp rings requires the treatment of 10 heavy atoms (i.e., atoms heavier than hydrogen and helium) and 10 H atoms, a Cp^*_2M complex involves (i) 20 carbon atoms in addition to the metal and (ii) many more degrees of geometrical freedom associated with rotations of the five methyl groups on each Cp* ligand. This increase in computational expense only becomes more severe if Cp rings substituted with alkyl groups larger than methyl are considered.

A popular means to accurately account for important steric contributions from large groups while minimizing their computational cost is the employ of a hybrid quantum mechanical/molecular mechanical (QM/MM) method.[35,36] Such a model divides the system of interest into two or more parts, with the part of primary interest, e.g., a region experiencing bond making or breaking, being treated with ab initio or density functional theory (DFT), while the surrounding supporting structure is treated classically with molecular mechanics. One two-layer QM/MM model widely used for organometallic and inorganic systems is the Integrated Molecular Orbital Molecular Mechanics (IMOMM) model of Morokuma and co-workers,[37,38] which has also been generalized to a two-layer QM/QM version (Integrated Molecular Orbital Molecular Orbital, IMOMO) and to the multilayer/multimethod ONIOM model.[39–44] The IMOMM or IMOMO energy is written as

$$E_{\mathrm{IMOM}x} = E^{\mathrm{small}}_{\mathrm{QM}} + \left(E^{\mathrm{large}}_{y\mathrm{M}} - E^{\mathrm{small}}_{y\mathrm{M}}\right) \quad (1)$$

where the total energy of the system can be thought of as a sum of a high-level energy for a small subsystem together with a correction for the effect of the remainder of the system computed at a less costly level of theory, either molecular mechanics ($x = y =$ "M" in eq 1) or a lower level of electronic structure theory ($x =$ "O", $y =$ "Q" in eq 1). In the case of a sandwich metallocene, Cp^*_2M, a logical choice for the small subsystem might be the metal and two Cp rings, Cp being the result of replacing each C–methyl bond in Cp* with a hydrogen atom (see Scheme 1).

The "mechanical embedding" implicit in the IMOMM energy expression can usefully account for steric influence,[45] and indeed

† Part of the Sason S. Shaik Festschriff.

* Corresponding author. E-mail: cramer@umn.edu

‡ Taken in part from Lewin, J. L. Ph.D. Thesis, University of Minnesota, Minneapolis, MN, 2007.

10.1021/jp711830a CCC: $40.75
Published on Web 06/21/2008

SCHEME 1

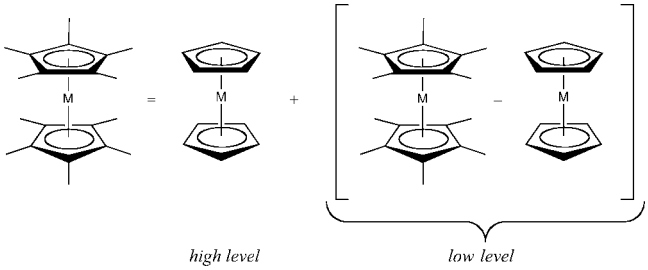

TABLE 1: CEP-31G Effective Core Potential Parameters for B, C, N, O, and F[a]

atom	ECP	A_{lk}	n_{lk}	B_{lk}
B	V_p	−0.88355	1	5.64622
	V_{s-p}	1.89572	0	1.92061
		10.97543	2	5.55177
C	V_p	−0.89371	1	8.56468
	V_{s-p}	1.92926	0	2.81497
		14.88199	2	8.11296
N	V_p	−0.91212	1	11.99686
	V_{s-p}	1.93565	0	3.83895
		21.73355	2	11.73247
O	V_p	−0.92550	1	16.11718
	V_{s-p}	1.96069	0	5.05348
		29.13442	2	15.95333
F	V_p	−0.93258	1	20.73959
	V_{s-p}	2.03649	0	6.60488
		27.86279	2	18.24092

[a] Taken from ref 60.

Perrin et al.[18] found the partitioning in Scheme 1 to provide generally good agreement with fully optimized QM calculations for certain metallocene structures and reactions. However, as noted by Perrin et al., such performance cannot always be assured, particularly when modeling systems that feature a change in oxidation state or coordination number of the metal center, where electronic and steric properties of ligands are both important. Thus, for instance, the IMOMM protocol of Scheme 1 cannot a priori account for the difference in gas-phase electron affinities of the Cp and Cp* radicals (1.29 eV for Cp* and 1.79 eV for Cp[46]) or the difference in 1-electron standard reduction potentials for the ferrocene/ferrocenium and decamethylferrocene/decamethylferrocenium couple in acetonitrile (590 mV),[47] which derives from the greater electron-donating character of Cp* compared to Cp, since the QM portion of the calculation never addresses a system larger than Cp. Another interesting case where the IMOMM formalism might be expected to fail is in predicting the effects of increasing Cp ring substitution on carbonyl stretching frequencies when such ligands are present in the metallocene. Zachmanoglou et al.[48] studied this effect on $\nu_{CO(avg)}$ for a series of $(Cp^R)_2Zr(CO)_2$ zirconocene complexes and observed a noticeable decrease in the average stretching frequency as the degree of total $(Cp^R)_2$ methylation was increased.

In this work, we examine whether the electronic influence of Cp alkyl substituents may be captured within the simplicity of the IMOMM formalism by modifying the electronic character of the Cp-ring carbon atoms. In particular, we develop a modified effective core potential (ECP) for these C atoms to be used in the QM portion of the calculation. Such an approach has some precedent in the work of Zhang et al.,[49] who modified a fluorine ECP so that the resulting atom mimicked a methyl group (for use as a link atom in protein modeling at the QM/MM level). More globally, Alary and co-workers have described methods for replacing entire ligands in organometallic complexes by ad hoc construction of optimized effective group potentials.[50,51] As part of our work, we also examine the utility of combining a modified pseudopotential with the IMOMO approach. While the IMOMO method *does* treat the full system at a quantum mechanical level, the use of a very low level of electronic structure theory may still require some adjustment in the high-level calculation in order to maximize accuracy.

Theoretical Methods. ***Electronic Structure Calculations.*** All wave function theory (WFT) and DFT calculations employed the Gaussian03 suite of programs, including its implementation of IMOMM and IMOMO.[52] Density functionals used in this work include the hybrid B3LYP functional[53–56] and generalized gradient approximation *m*PW functional.[57–59]

The CEP-31G(d) effective core potentials (see below) and associated basis functions of Stevens et al.[60] were used for metal atoms, carbon, and oxygen, and the Pople 6-31G basis for hydrogen;[61] we hereafter refer to this basis set combination as DZP. When modified ECPs are used on ring carbons, the CEP-31G(d) valence basis functions are used without modification. IMOMM calculations used the universal force field[62] (UFF) for the MM level of theory, while IMOMO calculations used ab initio Hartree−Fock (HF) theory[63] with the minimal STO-3G basis set[61] for the low-level MO theory; these combinations of IMOMM and IMOMO are at times referred to in the text as DFT/DZP:UFF and DFT/DZP:HF/STO-3G, respectively.

Pseudopotentials. The form of the CEP-31G carbon pseudopotential is[60]

$$r^2V(r) = V_p(r) + V_{s-p}(r)|Y_0^0\rangle\langle Y_0^0| \tag{2}$$

where r is the radial distance from the nucleus, Y_0^0 is the s-type spherical harmonic, and the remaining potential functions are expanded as sums of gaussians

$$r^2V(r) = \sum_k A_{lk} r^{n_{lk}} e^{-B_{lk}r^2} \tag{3}$$

where V is the total effective potential, r is the distance from the nucleus, l is the angular momentum quantum number, k is an index, n_{lk} is equal to 0, 1, or 2, and A_{lk} and B_{lk} are parameters. The complete expansion for the 1s core of carbon is expressed as

$$r^2V(r) = \underbrace{A_{11}r^1e^{-B_{11}r^2}}_{\text{p potential}} + \underbrace{\left(A_{01}r^0e^{-B_{01}r^2} + A_{02}r^2e^{-B_{02}r^2}\right)}_{\text{s–p potential}} \tag{4}$$

where A_{01}, A_{02}, A_{11}, B_{01}, B_{02}, and B_{11} are adjustable parameters. Table 1 provides the parameter values defined for B, C, N, O, and F in the CEP-31G ECP.

Parameter optimizations for modified C pseudopotentials made use of a microgenetic algorithm[64] (μGA) that optimizes a given set of parameters by maximizing a fitness function, which we chose to be the negative of the root-mean-square error (RMSE) in properties predicted at the IMOMM or IMOMO levels compared to fully quantum mechanical B3LYP/DZP results. We thus maximized

$$Q = -\sqrt{\frac{\sum_{i,j}\left[w_i\left(q^i_{j,(\text{calc})} - q^i_{j,(\text{ref})}\right)\right]^2}{N}} \tag{5}$$

where Q is the unitless fitness function, w_i is the weighting factor of the property i, $q^i_{j,(\text{calc})}$ is the value of the IMOMM or IMOMO

TABLE 2: Data for Different IMOMM + ECP Combinations Compared to B3LYP/DZP Reference Data

property	Cp*M	QM[a]	UFF[b]+C	UFF+C_N	UFF+C_O	UFF+C_{MM}
C−C dist (Å)	Li	1.448	1.448	1.433	1.428	1.423
	Na	1.448	1.449	1.433	1.428	1.421
	K	1.445	1.447	1.431	1.426	1.417
C−M dist (Å)	Li	2.130	2.163	2.152	2.145	2.138
	Na	2.578	2.584	2.575	2.569	2.557
	K	3.053	3.041	3.035	3.031	3.015
dipole moment (D)	Li	3.4	4.2	3.9	3.8	3.7
	Na	6.8	7.2	6.9	6.7	6.5
	K	10.5	10.4	10.0	9.8	9.7
vertical IP (eV)	Li	6.6	7.8	7.5	7.3	7.2
	Na	5.8	6.9	6.6	6.4	6.3
	K	5.2	6.1	5.8	5.7	5.6
binding energy (kcal mol^{-1})	Li	−165.2	−132.7	−135.9	−138.7	−145.5
	Na	−129.7	−105.1	−108.0	−110.2	−114.2
	K	−105.6	−84.2	−86.4	−88.1	−90.9
M-ring stretch (cm^{-1})	Li	544.2	544.2	552.9	562.4	589.9
	Na	285.6	321.8	322.8	325.2	343.0
	K	227.5	281.7	283.8	285.4	299.3
fitness (−RMSE)		0.00	−50.99	−44.67	−40.59	−38.14

[a] B3LYP/DZP calculation on complete Cp* systems. [b] IMOMM with B3LYP/DZP as high-level theory applied to Cp subsystem and UFF as low-level theory.

computed property i at occurrence j, $q^i_{j,(\mathrm{ref})}$ is the value of the property i at occurrence j computed at the fully optimized QM B3LYP/DZP reference level, and N is the number of total data.

In the μGA runs, each generation had a population of eight individuals (i.e., parameter sets) all having six chromosomes (ECP parameters). Elitism was enforced; thus, the fittest individual from each generation, as determined by Q, was retained in the subsequent generation. For the C_{HF} ECP, the optimization was found to converge (no change in the sixth decimal place of Q) after three independent μGA runs of 262, 217, and 176 generations. The optimization of the C_{MM} ECP was terminated after 9 independent or semi-independent (some were sequential) μGA runs totalling 884 generations. Multiple runs starting from different initial parameter values were undertaken to ensure adequate coverage of parameter space.

Results and Discussion

Initial Parameter Optimization. To begin, the structures of free Cp^{*-}, Cp*Li, Cp*Na, Cp*K, and the radical cations of the three Cp*M complexes were optimized at the fully QM B3LYP/DZP level. This choice of systems was motivated in part because of the efficiency in using half-sandwich complexes and in part because of the span of covalent to ionic character that should be associated with moving from Li^+ to Na^+ to K^+ as the gas-phase counterion. The Cp^{*-} anion and the neutral half-sandwich complexes have geometries belonging to the C_5 point group; Jahn−Teller distortion reduces the symmetry of the radical cations to C_s. The frozen B3LYP/DZP radical cation geometries were used to evaluate vertical ionization potentials (eV, from the neutral) with different pseudopotential parameters. In the remaining cases, full geometry reoptimizations were performed for all parameter choices and ring C−C distances (Å, one unique value by symmetry), ring C−metal distances (Å, one unique value by symmetry), dipole moments (D), metal binding energies (kcal mol^{-1}, computed as the energy difference between the half-sandwich complex and the separated alkali metal cation and Cp^{*-}), and metal−ring stretching frequencies (cm^{-1}, this mode was easily identified in every case) were compared to fully QM values in construction of the fitness function. The weighting factors were chosen as 100.0 $Å^{-1}$ for bond lengths, 10.0 D^{-1} for dipole moments, 10.0 eV^{-1} for ionization potentials, 1.0 mol $kcal^{-1}$ for binding energy, and 0.2 cm for stretching frequencies. These choices are perforce arbitrary, but serve to "chemically normalize" the allowed variations in the target properties. Tables 2 and 3 provide the target data, the data computed with different pseudopotential parameters in the IMOMM and IMOMO approximations, and the associated fitness functions. Specifically, the tables provide results for the normal C 1s core, cores in which the ECPs of N and O are used in place of the normal C core, referred to respectively as C_N and C_O, and the ECPs determined from μGA optimization, designated C_{MM} for the IMOMM case and C_{HF} for the IMOMO case. The optimized pseudopotential parameters for the IMOMM and IMOMO implementations are provided in Table 4. Choosing to consider the C_N and C_O protocols is clearly arbitrary, but the pseudopotential parameters are physically balanced for other first-row atoms, and consideration of their associated performance for the test-set data permits us to assess the sensitivity of the error function to variations in parameter space not associated with μGA optimization.

We consider the IMOMO data in Table 3 first. For both C−C and C−M distances, IMOMO with the unmodified carbon ECP gives distances slightly larger than the fully QM reference values; use of C_O improves C−M distances, and use of C_N improves C−C distances, though in each case the C−C distances are shorter than the QM reference. Dipole moments are moderately sensitive to the different ECPs, as are vertical ionization potentials, although no single ECP is optimal for both of these properties across all three complexes. Binding energies for unmodified carbon and C_N are predicted to be similar for all three complexes and agree reasonably well with reference values except for Cp*Li, where C_O gives much better agreement. Finally, M-ring stretching frequencies show similar behavior, but for this property the C_O ECP gives the best overall agreement with the B3LYP/DZP reference values. In the absence of any clear trends, it is unsurprising that the fitness function Q fails to vary much across unmodified C, C_N, and C_O; given the necessarily arbitrary nature of the fitness function, such small variation might even be regarded as insignificant.

TABLE 3: Data for Different IMOMO + ECP Combinations Compared to B3LYP/DZP Reference Data

property	Cp*M	QM[a]	HF[b] + C	HF + C_N	HF + C_O	HF + C_{HF}
C−C dist (Å)	Li	1.448	1.454	1.440	1.436	1.428
	Na	1.448	1.452	1.437	1.432	1.424
	K	1.445	1.450	1.434	1.429	1.421
C−M dist (Å)	Li	2.130	2.152	2.143	2.136	2.137
	Na	2.578	2.598	2.585	2.581	2.577
	K	3.053	3.102	3.092	3.085	3.074
dipole moment (D)	Li	3.4	3.6	3.3	3.1	3.3
	Na	6.8	6.8	6.4	6.2	6.5
	K	10.5	10.5	10.2	9.9	10.1
vertical IP (eV)	Li	6.6	6.8	6.6	6.3	6.6
	Na	5.8	6.4	6.4	5.7	6.4
	K	5.2	5.6	5.3	5.4	5.3
binding energy (kcal mol^{-1})	Li	−165.2	−158.5	−158.7	−165.0	−160.4
	Na	−129.7	−132.5	−133.2	−138.4	−134.3
	K	−105.6	−106.7	−106.1	−111.6	−106.8
M-ring stretch (cm^{-1})	Li	544.2	522.1	526.4	535.8	548.5
	Na	285.6	277.8	279.9	281.7	290.7
	K	227.5	222.6	222.4	223.6	228.9
fitness (−RMSE)		0.00	−13.02	−11.97	−14.66	−11.05

[a] B3LYP/DZP calculation on complete Cp* systems. [b] IMOMO with B3LYP/DZP as high-level theory applied to Cp subsystem and HF/STO-3G as low-level theory.

TABLE 4: Effective Core Potential Parameters Optimized for C in IMOMM and IMOMO Metallocene Calculations

atom	ECP	A_{lk}	n_{lk}	B_{lk}
C_{MM}	V_p	−0.18385	1	21.57558
	V_{s-p}	1.23783	0	6.91011
		35.75519	2	21.08175
C_{HF}	V_p	−0.79890	1	18.83507
	V_{s-p}	1.28382	0	2.47945
		14.31509	2	18.30750

The rightmost column in Table 3 shows the property values computed with the μGA optimized C_{HF} ECP. Overall this parametrized ECP exhibits marginally improved agreement with fully QM values, a notable exception being that ring C−C distances are predicted to be about 0.02 Å too short; this error appears to be balanced by improved predictions of C−M bond distances. Compared with the unmodified carbon, C_N, and C_O ECPs, M-ring stretching frequencies arguably show the greatest improvement compared with other properties. Consideration of the fitness function indicates that, relative to the unmodified C ECP, RMSE is reduced by about 15%. The small magnitude of this error reduction may possibly be attributable to the intrinsically good performance of the IMOMO model, even with HF/STO-3G for the low level. Prior to considering this point in more detail, we first examine the IMOMM results.

Similar to IMOMO, IMOMM (Table 2) gives acceptable agreement with C−C and C−metal distances, performing best with unmodified carbon ECPs for C−C distances and becoming shorter with C_N and C_O ECPs. Dipole moments for Cp*Li and Cp*Na show moderate to large deviations with unmodified C ECPs since within the IMOMM formalism the dipole moment is entirely determined from the QM calculation on the CpM subsystem. In the case of Cp*K, the complex is sufficiently ionic that this approximation does not seem to have as large an impact. The dipole moments decrease in magnitude along the ECP series C > C_N > C_O. Vertical ionization potentials also show a decreasing trend along this series, going from a deviation of about 1 eV with unmodified C (again owing to no contribution from QM methyl groups) to about 0.6 eV for C_O. Binding energies are poorly reproduced with IMOMM, being underestimated in every case, but trending slightly toward the fully QM values on going from C to C_N to C_O. M-ring stretches show limited sensitivity to ECP in Cp*Na and Cp*K, and predictions are not particularly good. The fitness function Q now spans a larger range of RMSE across the 3 ECPs, about 20%, than was true for the IMOMO model. Optimization of the pseudopotential parameters using the μGA gives a fitness function slightly improved by comparison to the C_O ECP. The greatest improvement is in binding energies, but these remain significantly underestimated. Moreover, inspection of the parameter values in Table 4 indicates that they have moved rather far from the ranges sampled by the first-row atoms (cf. Table 1), suggesting that the maximum in Q for C_{MM} may not be particularly physical. To evaluate this point more closely, we now consider a test metallocene system showing spectroscopic variations associated with varying methyl substitution.

Carbonyl Stretching Frequencies in $(Cp^R)_2Zr(CO)_2$. In an interesting spectroscopic study of metallocene carbonyls, Zachmanoglou et al.[48] observed that the average carbonyl stretching frequency $\nu_{CO(avg)}$ decreased by about 3.2 cm^{-1} per methyl substituent in a series of 10 different zirconocene carbonyls varying in the degree of methylation of the Cp rings. Such a trend is consistent with increased electron donation to the metal being partially delocalized into the carbonyl π^* acceptor orbitals. Thus, on going from $Cp_2Zr(CO)_2$ to $Cp^*_2Zr(CO)_2$, a total decrease in $\nu_{CO(avg)}$ of 32 cm^{-1} was observed with changes roughly linear in number of methyl groups (irrespective of positional substitution). The experimental data, together with data computed at the IMOMO level using different ECPs, are provided in Table 5.

Fully QM calculations give a total unsigned error (ΣUE) of 22 cm^{-1} after application of a scale factor of 0.9716. In this case, the scale factor has been optimized to minimize ΣUE for this specific problem, but it is very typical of scale factors optimized for the B3LYP functional with similar basis sets and broader compilations of experimental frequencies.[63] The slope of the line fit to the fully QM calculations is −2.69, compared to −3.34 for the experimental data, but it should be noted that the slope is rather sensitive to points at the ends of the range and moreover that there is not any reason to expect a perfect linear relationship in any case, so ΣUE is arguably a more relevant measure of accuracy.

TABLE 5: Experimental and Computed $\nu_{CO(avg)}$ Values (cm^{-1}) for $(Cp^R)_2Zr(CO)_2$ complexes at B3LYP and IMOMO/B3LYP/HF Levels

complex	no. Me	exptl[a]	QM[b]	C_B[c]	C	C_N	C_O	C_F	C_{HF}
Cp_2Zr	0	1932	1934	1934	1934	1934	1934	1934	1934
$Cp(Cp^{Me})Zr$	1	1930	1928	1936	1932	1929	1927	1927	1929
$(Cp^{Me})_2Zr$	2	1927	1922	1939	1929	1923	1921	1920	1923
$(Cp^{1,2\text{-}Me_2})_2Zr$	4	1922	1919	1944	1926	1925	1925	1923	1927
$Cp(Cp^{Me_4})Zr$	4	1919	1919	1937	1930	1924	1921	1919	1923
$Cp(Cp^*)Zr$	5	1920	1920	1941	1927	1920	1914	1912	1918
$(Cp^{1,2,4\text{-}Me_3})_2Zr$	6	1913	1913	1944	1924	1913	1908	1904	1912
$(Cp^{Me})(Cp^*)Zr$	6	1913	1914	1944	1925	1914	1908	1904	1912
$(Cp^{Me_4})_2Zr$	8	1905	1909	1948	1922	1909	1901	1897	1906
Cp^*_2Zr	10	1900	1904	1948	1919	1903	1895	1889	1901
ΣUE[c]			22	236	85	24	41	57	21
slope[d]		−3.34	−2.69	1.47	−1.39	−2.89	−3.83	−4.42	−3.25
R^2 [d]		0.979	0.952	0.849	0.936	0.937	0.929	0.943	0.926

[a] Ref 48. [b] Fully QM B3LYP/DZP frequencies scaled by 0.9716. [c] IMOMO with B3LYP/DZP as high-level theory applied to Cp subsystem and HF/STO-3G as low-level theory; frequencies scaled by 0.9716. Sum of unsigned error over all 10 predictions. [d] From linear regression of theoretical data on experimental data.

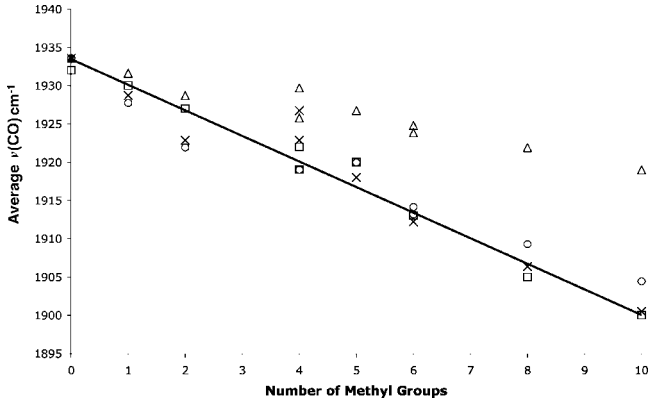

Figure 1. Experimental data (squares) and predictions from fully QM B3LYP/DZP (circles) and IMOMO calculations using either CEP C (triangles) or optimized C_{HF} (crosses) carbon 1s pseudopotentials. A best-fit line to the experimental data is provided to guide the eye, but there is no requirement that a linear relationship exist.

Employing the IMOMO approach with various choices for the C 1s pseudopotential leads to varying results. Using the B core, for instance, gives a very large ΣUE of 236 cm^{-1}, and a positive slope is predicted as a function of methyl substitution instead of a negative one. With the original CEP C 1s core potential, the slope has the right sign, but it is fairly small (−1.39) with $\Sigma UE = 85$ cm^{-1}. It is noteworthy that this poor performance is not associated with use of the scale factor chosen based on the fully QM calculations. Even if a scale factor is optimized for the IMOMO calculation with the CEP C 1s pseudopotential equal to 0.9682, which is quite close to the 0.9716 value for the fully QM case, the ΣUE value remains a sizable 54 cm^{-1}. The lowest ΣUE value, 21 cm^{-1}, is obtained with the C_{HF} pseudopotential, whose parameters were optimized for the alkali metallocenes above and applied without adjustment to this problem. Figure 1 illustrates the performance of three computational models compared to experiment. The use of CEP N parameters also gives good predictions, reflecting the similarity between the parameters in these two pseudopotentials.

As a test of transferability, we also computed the average carbonyl stretching frequencies in the zirconocene series at the *m*PW/DZP level of theory. The results, in Table 6, indicate that the IMOMO model with C_{HF} carbon 1s pseudopotentials improves on the fully QM prediction by about 25%, based on comparison of ΣUE values. Such improvement is purely fortuitous, since the pseudopotential was parametrized to reproduce fully QM predicted values and not experimental values, but it indicates its general applicability to this problem.

TABLE 6: Experimental and Computed $\nu_{CO(avg)}$ Values (cm^{-1}) for $(Cp^R)_2Zr(CO)_2$ Complexes at *m*PW and IMOMO/*m*PW/HF Levels

complex	no. Me	exptl[a]	QM[b]	C_{HF}
Cp_2Zr	0	1932	1935	1935
$Cp(Cp^{Me})Zr$	1	1930	1929	1929
$(Cp^{Me})_2Zr$	2	1927	1923	1924
$(Cp^{1,2\text{-}Me_2})_2Zr$	4	1922	1916	1925
$Cp(Cp^{Me_4})Zr$	4	1919	1926	1925
$Cp(Cp^*)Zr$	5	1920	1920	1921
$(Cp^{1,2,4\text{-}Me_3})_2Zr$	6	1913	1912	1913
$(Cp^{Me})(Cp^*)Zr$	6	1913	1914	1914
$(Cp^{Me_4})_2Zr$	8	1905	1909	1907
Cp^*_2Zr	10	1900	1903	1902
ΣUE[c]			30	22
slope[d]		−3.34	−3.02	−3.23
R^2 [d]		0.979	0.904	0.936

[a] Ref 48. [b] Fully QM *m*PW/DZP frequencies scaled by 1.0068. [c] IMOMO with B3LYP/DZP as high-level theory applied to Cp subsystem and HF/STO-3G as low-level theory; frequencies scaled by 1.0068. Sum of unsigned error over all 10 predictions. [d] From linear regression of theoretical data on experimental data.

Application of the IMOMM protocol, instead of the IMOMO one, to the zirconocene test set delivered less impressive results. As shown in Table 7, use of the default CEP C pseudopotential permits essentially no discrimination between the various methylated species, as might be expected given that the only manifestation of the methyl groups is steric during the UFF portion of the calculations. As methylated positions are represented with C_N and C_O pseudopotentials, on the other hand, some limited ability to capture the methylation trend is exhibited. In the case of C_O, the ΣUE value is 60 cm^{-1} compared to the fully QM value of 22 cm^{-1}, and the per-methyl effect is underestimated by about 33%. Use of the C_{MM} pseudopotential led to very poor results. In addition to the large ΣUE value, inspection of the various zirconocene geometries showed that in many cases the substituted Cp rings included significantly pyramidalized carbon atoms, i.e., the geometries were significantly distorted from those predicted at the fully QM level.

By consideration of this observation and recognition of the similar fitness function values for C_O and C_{MM} in Table 2, it would appear that μGA optimization of the pseudopotential parameters for the IMOMM model wandered into an unphysical

TABLE 7: Experimental and Computed $\nu_{CO(avg)}$ Values (cm^{-1}) for $(Cp^R)_2Zr(CO)_2$ Complexes at B3LYP and IMOMM/B3LYP/UFF Levels

complex	no. Me	exptl[a]	QM[b]	C[c]	C_N	C_O	C_{MM}
Cp_2Zr	0	1932	1934	1934	1934	1934	1934
$Cp(Cp^{Me})Zr$	1	1930	1928	1935	1932	1931	1930
$(Cp^{Me})_2Zr$	2	1927	1922	1936	1931	1928	1927
$(Cp^{1,2\text{-}Me_2})_2Zr$	4	1922	1919	1936	1928	1923	1921
$Cp(Cp^{Me_4})Zr$	4	1919	1919	1935	1929	1927	1925
Cp(Cp*)Zr	5	1920	1920	1934	1928	1927	1947
$(Cp^{1,2,4\text{-}Me_3})_2Zr$	6	1913	1913	1936	1926	1921	1919
(Cp^{Me})(Cp*)Zr	6	1913	1914	1936	1926	1924	1941
$(Cp^{Me_4})_2Zr$	8	1905	1909	1935	1923	1919	1936
Cp^*_2Zr	10	1900	1904	1934	1918	1909	1936
ΣUE[c]			22	166	92	60	138
slope[d]		−3.34	−2.69	[e]	−1.43	−2.07	[e]
R^2 [d]		0.979	0.952	0.021	0.970	0.884	0.063

[a] Ref 48. [b] Fully QM B3LYP/DZP frequencies scaled by 0.9716. [c] IMOMM with B3LYP/DZP as high-level theory applied to Cp subsystem and UFF as low-level theory; frequencies scaled by 0.9716. Sum of unsigned error over all 10 predictions. [d] From linear regression of theoretical data on experimental data. [e] Not statistically meaningful.

region of parameter space in order to achieve very small improvement in Q, possibly because symmetry constraints in the parametrization protocol masked geometric errors that might otherwise have prevented such evolution. In any case, noting that the C_O pseudopotential parameters provide a nearly optimal value of Q and that they are "physical" in the sense that they have been optimized for an actual first-row atom, they probably constitute a best choice for similar situations, although the performance is only marginally satisfactory.

Given the reasonable, and more physical, behavior of the IMOMO with C_{HF} model, we close by noting that, compared to a full QM calculation (B3LYP/DZP) on $Cp^*_2Zr(CO)_2$, IMOMO (B3LYP/DZP:HF/STO-3G) with C_{HF} ECPs provides a 2-fold increase in speed for geometry optimizations and a 5-fold increase in speed for frequency calculations on four 2.6 GHz AMD Opteron processors with 7 GB of memory. Such speed increases may be useful for in silico catalyst design where surveys of many different substituted Cp rings may be envisioned.

Acknowledgment. This work was supported by the U.S. National Science Foundation (CHE-0610183). We thank Ben Lynch for assistance with implementation of the μGA.

References and Notes

(1) Long, N. J. Metallocenes. In *An Introduction to Sandwich Complexes*; Blackwell Science: Malden, MA, 1998.
(2) Davies, J. A.; Watson, P. L.; Liebman, J. F.; Greenberg, A. *Selective Hydrocarbon Activation: Principles and Progress*; VCH Publishers, Inc.: New York, 1990.
(3) Kaminsky, W.; Arndt, M. *Adv. Polym. Sci.* **1997**, *127*, 143.
(4) Chirik, P. J.; Bercaw, J. E. In *Metallocenes: Synthesis, Reactivity, Applications*; Togni, A., Halterman, R. L., Eds.; Wiley-VCH: Weinheim, 1998; Vol. 1, Chapter 3.
(5) Molander, G. A. *Chemtracts—Org. Chem.* **1998**, *11*, 237–263.
(6) Watson, P. L.; Parshall, G. W. *Acc. Chem. Res.* **1985**, *18*, 51–56.
(7) Watson, P. L. In *Selective Hydrocarbon Activation: Principles and Progress*; Davies, J. A., Watson, P. L., Greenberg, A., Liebman, J. F., Eds.; VCH: New York, 1990; pp 79−110..
(8) Crabtree, R. H. *The Organometallic Chemistry of the Transition Metals*, 3rd ed.; Wiley: New York, 2001.
(9) Folga, E.; Ziegler, T.; Fan, L. *New J. Chem.* **1991**, *15*, 741–748.
(10) Folga, E.; Ziegler, T. *Can. J. Chem.* **1992**, *70*, 333–342.
(11) Rappé, A. K.; Upton, T. H. *J. Am. Chem. Soc.* **1992**, *114*, 7507–7517.
(12) Deelman, B.-J.; Teuben, J. H.; Macgregor, S. A.; Eisenstein, O. *New J. Chem.* **1995**, *19*, 691–698.
(13) Maron, L.; Eisenstein, O. *J. Am. Chem. Soc.* **2001**, *123*, 1036–1039.
(14) Maron, L.; Eisenstein, O.; Alary, F.; Poteau, R *J. Phys. Chem. A* **2002**, *106*, 1797–1801.
(15) Maron, L.; Perrin, L.; Eisenstein, O. *J. Chem. Soc., Dalton Trans.* **2002**, 534–539.
(16) Perrin, L.; Maron, L.; Eisenstein, O. *Inorg. Chem.* **2002**, *41*, 4355–4362.
(17) Maron, L.; Perrin, L.; Eisenstein, O. *Dalton Trans.* **2003**, 4313–4318.
(18) Perrin, L.; Maron, L.; Eisenstein, O. *New J. Chem.* **2004**, *28*, 1255–1259.
(19) Barros, N.; Eisenstein, O.; Maron, L. *Dalton Trans.* **2006**, 3052–3057.
(20) Barros, N.; Eisenstein, O.; Maron, L.; Tilley, T. D. *Organometallics* **2006**, *25*, 5699–5708.
(21) Perrin, L.; Eisenstein, O.; Maron, L. *New J. Chem.* **2007**, *31*, 549–555.
(22) Werkema, E. L.; Maron, L.; Eisenstein, O.; Andersen, R. A. *J. Am. Chem. Soc.* **2007**, *129*, 2529–2541.
(23) Shilov, A. E.; Shul'pin, G. B. *Chem. Rev.* **1997**, *97*, 2879–2932.
(24) Fey, N. *J. Chem. Technol. Biotechnol.* **1999**, *74*, 852–862.
(25) Niu, S.; Hall, M. B. *Chem. Rev.* **2000**, *100*, 353–405.
(26) Rotzinger, F. P. *Chem. Rev.* **2005**, *105*, 2003–2038.
(27) Adamo, C.; Barone, V. *J. Comput. Chem.* **2000**, *21*, 1153–1166.
(28) Sherer, E. C.; Cramer, C. J. *Organometallics* **2003**, *22*, 1682–1689.
(29) Lewin, J. L.; Woodrum, N. L.; Cramer, C. J. *Organometallics* **2006**, *25*, 5906–5912.
(30) Woodrum, N. L.; Cramer, C. J. *Organometallics* **2006**, *25*, 68–73.
(31) Baisch, U.; Zeuner, M.; Barros, N.; Maron, L.; Schnick, W. *Chem.—Eur. J.* **2006**, *12*, 4785–4798.
(32) Roger, M.; Barros, N.; Arliguie, T.; Thuery, P.; Maron, L.; Ephritikhine, M. *J. Am. Chem. Soc.* **2006**, *128*, 8790.
(33) de Almeida, K. J.; Cesar, A. *Organometallics* **2006**, *25*, 3407–3416.
(34) Raynaud, C.; Perrin, L.; Maron, L. *Organometallics* **2006**, *25*, 3143–3151.
(35) *Combined Quantum Mechanical and Molecular Mechanical Models*; Gao, J., Thompson, M. A., Eds.; American Chemical Society: Washington, DC, 1998; Vol. 712.
(36) Senn, H. M.; Thiel, W. *Top. Curr. Chem.* **2007**, *268*, 173–290.
(37) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170–1179.
(38) Matsubara, T.; Maseras, F.; Koga, N.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 2573–2580.
(39) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959–1967.
(40) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357–19363.
(41) Svensson, M.; Humbel, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 3654–3661.
(42) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419–1432.
(43) Vreven, T.; Morokuma, K.; Farkas, Ö.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760–769.
(44) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theor. Comput.* **2006**, *2*, 815–826.
(45) Maseras, F. *Chem. Commun.* **2000**, 1821–1827.
(46) Bartmess, J. E. National Institute of Standards and Technology Webbook, webbook.nist.gov.
(47) Connelly, N. G.; Geiger, W. E. *Chem. Rev.* **1996**, *96*, 877–910.
(48) Zachmanoglou, C. E.; Docrat, A.; Bridgewater, B. M.; Parkin, G.; Brandow, C. G.; Bercaw, J. E.; Jardine, C. N.; Lyall, M.; Green, J. C.; Keister, J. B. *J. Am. Chem. Soc.* **2002**, *124*, 9525–9546.
(49) Zhang, Y. K.; Lee, T. S.; Yang, W. T. *J. Chem. Phys.* **1999**, *110*, 46–54.
(50) Alary, F.; Poteau, R.; Heully, J.; Barthelat, J.; Daudey, J. *Theor. Chem. Acc.* **2000**, *104*, 174–178.
(51) Carissan, Y.; Bessac, F.; Alary, F.; Heully, J.-L.; Poteau, R. *Int. J. Quantum Chem.* **2006**, *106*, 727–733.
(52) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz,

P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Pittsburgh, PA, 2003.

(53) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(54) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(55) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(56) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(57) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664–675.

(58) Perdew, J. P. In *Electronic Structure of Solids '91*; Ziesche, P., Eschrig, H., Eds.; Akademie Verlag: Berlin, 1991, pp 11−20.

(59) Perdew, J.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244–13249.

(60) Stevens, W.; Basch, H.; Krauss, M. *J. Chem. Phys.* **1984**, *81*, 6026.

(61) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.

(62) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.

(63) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; John Wiley & Sons: Chichester, 2004.

(64) Carroll, D. L. *AIAA J.* **1996**, *34*, 338–346.

JP711830A