Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.


# Ionization energies of metallocenes: A coupled cluster study of cobaltocene

## Electronic Supplementary Information (ESI)

Heiðar Már Aðalsteinsson, Ragnar Bjornsson[a],

E-mail of Corresponding author:
[a]ragnar.bjornsson@cea.fr

# Contents

# S.1 Geometries

Geometries of all metallocenes from section **3.4** in the main article with both BP86 and PBE0 are available in the ESI as xyz files in Cartesian coordinates (Å). All geometries were optimized with D3BJ dispersion and def2-TZVP basis sets. Also included are optimized geometries of $Cp_2Mn$ ($^6A_{1g}$) with D4 dispersion.

# S.2 Cobaltocene MO diagrams

Molecular orbital diagrams for $Cp_2Co$ and $Cp_2Co^+$ were plotted for several DFT methods. Quasi-restricted orbitals (QROs) were used for neutral $Cp_2Co$ (open shell) and canonical orbitals for $Cp_2Co^+$ (closed-shell), which are the references used in the DLPNO-CCSD(T) calculations. The MO diagram for the cation was omitted from the main text due to its size, but both species are shown here.

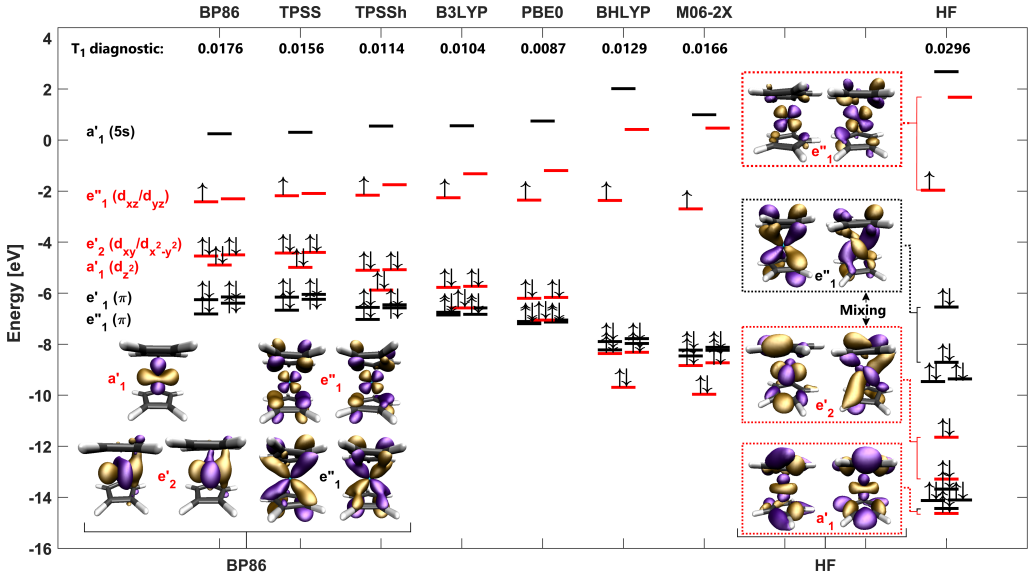


**Figure S.1** MO diagrams from QRO orbitals of neutral $Cp_2Co$ calculated at the cc-pVQZ-DK level. Methods are shown at the top with their corresponding $T_1$ diagnostic from CCSD calculations. Selected orbitals from BP86 and HF calculations were plotted and are labeled at the bottom. Cobalt d-orbitals are marked red.

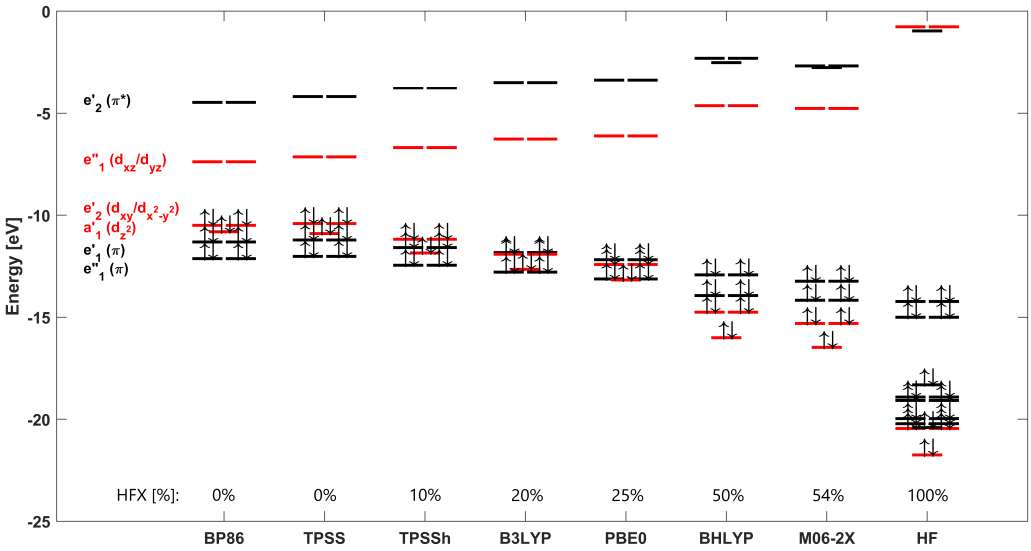


**Figure S.2** MO diagrams from canonical orbitals of the $Cp_2Co^+$ cation, calculated at the cc-pVQZ-DK level. The DFT method for each MO diagram is labeled at the bottom along with its corresponding fractional Hartree-Fock exchange (HFX). Cobalt d-orbitals are marked red.

As discussed in the main text, the ligand field splitting is increased with increasing HF exchange, eventually lowering the Co orbitals below the $\pi$ orbitals of the Cp rings. This effect is even more exaggerated for $Cp_2Co^+$ where B3LYP and all methods with higher HF exchange have Cp $\pi$ orbitals as their HOMO. Moreover, the Co 4s orbital is the LUMO when calculated with HF.

# S.3 Convergence of reference energies

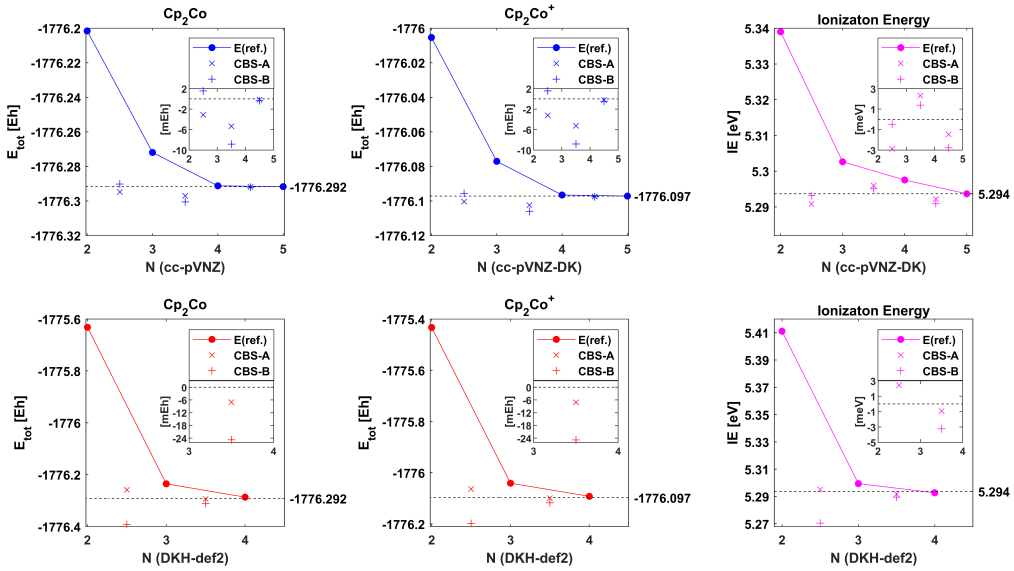


**Figure S.3** Convergence of reference energies from DLPNO-CCSD(T) calculations of $Cp_2Co$ (left) and $CpCo^+$ (middle) is shown with respect to basis set size, along with convergence of ionization energies (right), with both cc-pVnZ (top, blue) and def2 (bottom, red) basis sets. BP86 reference orbitals and the DKH Hamiltonian (and DKH basis sets) were used in all calculations. The cc-pV5Z-DK basis set is used as the CBS estimate.

Several ways to extrapolate reference energies have been proposed. Two methods were tested here, labeled CBS-A and CBS-B in Fig. S.3. The first one, CBS-A, is an expression from Karton and Martin,[1] shown in Eq. S.1, where X is the cardinal number of a given basis set and $E_{SCF}^{(X)}$ is the corresponding SCF energy. The constant $\alpha$ was optimized by extensive benchmarking for CBS(2/3) and CBS(3/4) SCF energies for several basis set families.[2] This constant varies between basis set families, and also between CBS(2/3) and CBS(3/4) for the same basis set family. The equation can be easily solved for A and $E_{SCF}^{(\infty)}$ by inserting the basis cardinal numbers and associated SCF energies.

$$E_{SCF}^{(X)} = E_{SCF}^{(\infty)} + Aexp(-\alpha\sqrt{X}) \tag{S.1}$$

$$E_{SCF}^{(\infty)} = \frac{X^{\beta}E_{SCF}^{(X)} - Y^{\beta}E_{SCF}^{(Y)}}{X^{\beta} - Y^{\beta}} \tag{S.2}$$

The other method, CBS-B, is based on the equation commonly used for extrapolation of correlation energies suggested by Helgaker and coworkers,[3] and is shown in Eq. S.2. Here X and Y are the basis set cardinal numbers, and $E_{SCF}^{(X)}$ and $E_{SCF}^{(Y)}$ are the corresponding SCF energies. This expression was used by Neese and coworkers in recent work with $\beta$ = 3.9 in SCF energy extrapolations.[4]

BP86 reference orbitals were used (QRO for open shell), and the DKH Hamiltonian was used throughout along with the corresponding basis sets (Fig. S.3). The CBS-A scheme performs better overall, however, CBS(3/4) extrapolations deviate 5 - 6 mEh from the reference (cc-pV5Z-DK) for both def2 and cc-pVNZ basis set families. The HF reference energies using non-selfconsistent BP86 orbitals seem to be close to convergence at the QZ level, resulting in the CBS(3/4) extrapolation overshooting the basis set limit. The CBS(4/5) extrapolation using cc-pV(Q/5)Z-DK basis are likely more reliable.

These are considerably larger errors than previously reported in studies of SCF energy extrapolation schemes, e.g. errors of 2 - 3 mEh for CBS(2/3) extrapolations and < 0.5 mEh for CBS(3/4) were reported by Neese and Valeev.[2] They also noted that extrapolations of SCF energies seem to be much more sensitive than those of correlation energies. As these extrapolation schemes were optimized through benchmarking with self-consistent HF energies, we believe the errors found in this work are likely due to the use of non-selfconsistent reference energies (i.e. HF energies determined on top of KS-DFT orbitals and orbital transformations (QRO) for open shell systems. Therefore, it was decided to use reference energies from the larger basis set and only extrapolate correlation energies in CBS extrapolations.

# S.4 Total metallocene energies

Here, we document the total energies associated with tables 3 and 5 from the main text. In Table S.1, we show the total energies for cobaltocene corresponding to the highest level of theory allowed by our computational resources from table 3 in the main text, i.e. cc-pwCV5Z energies for $E(ref.)$, CPS(2) CBS(4/5) energies for $E(CCSD)$ and $E(T_0)$ and CPS(2), CBS(3/4) energies for $E(T_1)_{corr}$ and $E(CV)_{corr}$, as described in Eq. 4 in the main text (see computational details).

Tables S.2 and S.3 show the total energies for the metallocene series from table 5 in the main text, calculated with protocol (1) and (2) using PBE0 reference orbitals, which corresponds to Eqs. 5 and 6 from the main text.

**Table S.1** Total energies for the highest level of theory from table 3 of the main text [a]

| | $E(ref.)$ | $E(CCSD)$ | $E(T_0)$ | $\Delta E(T_1)^b$ | $E(CV)^b$ | $E(ZPE)$ | $E(Tot)$ |
|---|---|---|---|---|---|---|---|
| | | | | Adiabatic IE | | | |
| $Cp_2Co$ | -1776.28810 | -2.90981 | -0.17212 | -0.02726 | -0.72662 | 0.16068 | -1779.96323 |
| $Cp_2Co^+$ | -1776.09358 | -2.90778 | -0.18058 | -0.02769 | -0.72642 | 0.16604 | -1779.77001 |
| | | | | Vertical IE | | | |
| $Cp_2Co^{+c}$ | -1776.08922 | -2.90082 | -0.18131 | -0.02774 | -0.72589 | - | -1779.92498 |

[a] All energies are reported in Hartree units [Eh]. [b] $\Delta E(T_1)$ and $E(CV)$ corrections were calculated as $E(T_1)$ - $E(T_0)$ and $E(AE)$ - $E(FC)$ at CPS(2) CBS(3/4) level. [c] Vertical IEs do not have ZPE corrections, so the ZPE correction should be subtracted from the total $Cp_2Co$ energy for calculation of the vertical IE.

**Table S.2** Total energies for metallocenes from protocol(1) with PBE0 reference orbitals (Section 3.4) [a]

| | $E(ref.)$ | $E(CCSD)$ | $E(T_0)$ | $\Delta E(T_1)^b$ | $G_{corr}$ | $E(Tot)$ |
|---|---|---|---|---|---|---|
| $Cp_2V$ | -1332.90851 | -2.39824 | -0.13534 | -0.00843 | 0.12451 | -1335.32601 |
| $Cp_2V^+$ | -1332.66643 | -2.38900 | -0.14363 | -0.01054 | 0.12611 | -1335.08349 |
| $Cp_2Cr$ | -1434.25456 | -2.52799 | -0.14893 | 0.00342 | 0.12416 | -1436.80390 |
| $Cp_2Cr^+$ | -1434.13714 | -2.45327 | -0.14455 | 0.00843 | 0.12606 | -1436.60047 |
| $Cp_2Mn$ | -1542.08011 | -2.45510 | -0.12588 | -0.02838 | 0.12057 | -1544.56890 |
| $Cp_2Mn^+$ | -1541.69690 | -2.57672 | -0.15582 | -0.03581 | 0.12629 | -1544.33896 |
| $Cp_2Fe$ | -1655.84668 | -2.75768 | -0.16837 | 0.00646 | 0.12984 | -1658.63643 |
| $Cp_2Fe^+$ | -1655.68955 | -2.67650 | -0.16151 | 0.01347 | 0.12797 | -1658.38612 |
| $Cp_2Co$ | -1776.41101 | -2.78743 | -0.16075 | -0.01100 | 0.12303 | -1779.24716 |
| $Cp_2Co^+$ | -1776.19592 | -2.80375 | -0.17109 | -0.01237 | 0.13122 | -1779.05191 |
| $Cp_2Ni$ | -1903.71035 | -2.79544 | -0.14777 | -0.01500 | 0.12204 | -1906.54652 |
| $Cp_2Ni^+$ | -1903.43167 | -2.82737 | -0.16318 | -0.01863 | 0.12375 | -1906.31710 |

[a] All energies are reported in Hartree units [Eh]. Energies were calculated according to Eq. 5 in the main text (Protocol (1)). [b] $\Delta E(T_1)$ is calculated as $E(T_1)$-$E(T_0)$ with N.PNO cut-offs in CBS(3/4) extrapolations

**Table S.3** Total energies for metallocenes from protocol(2) with PBE0 reference orbitals (Section 3.4) [a]

| | $E(ref.)$ | $E(CCSD)$ | $E(T_0)$ | $\Delta E(T_1)^b$ | $G_{corr}$ | $E(Tot)$ |
|---|---|---|---|---|---|---|
| $Cp_2V$ | -1332.88801 | -2.25809 | -0.12327 | -0.00554 | 0.12451 | -1335.15040 |
| $Cp_2V^+$ | -1332.64616 | -2.24993 | -0.13205 | -0.00765 | 0.12611 | -1334.90968 |
| $Cp_2Cr$ | -1434.23406 | -2.38457 | -0.13655 | -0.01130 | 0.12416 | -1436.64232 |
| $Cp_2Cr^+$ | -1434.11690 | -2.31418 | -0.13294 | -0.00702 | 0.12606 | -1436.44498 |
| $Cp_2Mn$ | -1542.05912 | -2.31285 | -0.11322 | -0.00468 | 0.12057 | -1544.36930 |
| $Cp_2Mn^+$ | -1541.67661 | -2.43550 | -0.14382 | -0.01153 | 0.12629 | -1544.14117 |
| $Cp_2Fe$ | -1655.82584 | -2.61312 | -0.15567 | -0.02039 | 0.12984 | -1658.48518 |
| $Cp_2Fe^+$ | -1655.66916 | -2.53555 | -0.14959 | -0.01376 | 0.12797 | -1658.24009 |
| $Cp_2Co$ | -1776.39015 | -2.64171 | -0.14741 | -0.01758 | 0.12303 | -1779.07382 |
| $Cp_2Co^+$ | -1776.17538 | -2.66021 | -0.15881 | -0.01896 | 0.13122 | -1778.88214 |
| $Cp_2Ni$ | -1903.68948 | -2.65001 | -0.13463 | -0.01263 | 0.12204 | -1906.36471 |
| $Cp_2Ni^+$ | -1903.41119 | -2.68386 | -0.15038 | -0.01619 | 0.12375 | -1906.13787 |

[a] All energies are reported in Hartree units [Eh]. Energies were calculated according to Eq. 6 in the main text (Protocol (2)). [b] $\Delta E(T_1)$ is calculated as $E(T_1)$-$E(T_0)$ with $T_{CutPNO} = 10^{-6}$ and other cut-offs at N.PNO in M-CBS(3/4) extrapolations

# S.5 Reference determinant issue for canonical CCSD(T)

When comparing DLPNO-CCSD(T) to canonical CCSD(T) (Section 3.1) an unexpected effect was observed in the canonical case for the closed shell $Cp_2Co^+$. Significantly more triples correlation energy is recovered when performing unrestricted CCSD(T) calculations using quasi-restricted KS orbitals (from the unrestricted KS solution) compared to the restricted CCSD(T) calculations using the restricted KS orbitals. This effect was not observed when using a HF reference determinant and only for canonical CCSD(T) (not the DLPNO form). Table S.4 compares all components of the total energy of $Cp_2Co^+$ with canonical CCSD(T), using both open-shell and closed-shell reference determinants from both BP86 and HF reference orbitals.

**Table S.4** All components of the total energy of $Cp_2Co^+$, calculated with canonical CCSD(T) with the cc-pVDZ basis set using RKS/UKS (BP86) and RHF/UHF reference orbitals, including ionization energies for each method[a]

| | $E(DFT)$ $[Eh]$ | $E(ref.)$[b] $[Eh]$ | $E(CCSD)$ $[Eh]$ | $E(T)$ $[Eh]$ | $IE$ $[eV]$ |
|---|---|---|---|---|---|
| RKS | -1780.6430 | -1776.0053 | -2.1558 | -0.1406 | 4.988 |
| UKS (QRO) | -1780.6430 | -1776.0053 | -2.1558 | -0.1458 | 4.844 |
| RHF | - | -1776.1831 | -2.0156 | -0.1108 | 4.849 |
| UHF (QRO) | - | -1776.1831 | -2.0156 | -0.1108 | 4.849 |

[a] All energies are reported in Hartree units except ionization energy, wich is in eV. Energies of neutral $Cp_2Co$ are not shown, as this effect was only observed for the cation.[b] $E(ref.)$ is the reference HF energy using either KS or HF orbitals.

As shown in the table, the IE is lowered by 0.144 eV by the additional triples energy captured when using QRO reference orbitals from a UKS calculation. After detailed analysis it was found that the reason for the difference arises due to the inclusion of the 4th-order doubles-triples term[5] in the triples correlation energy that plays a role for non-selfconsistent HF references. Consulting with the ORCA developers it was revealed that the term is inconsistently applied for canonical RKS-based CCSD(T) in the ORCA source code. This problem will be fixed in an upcoming ORCA release. The wrong energy appears to only arise when canonical restricted CCSD(T) calculations are performed using non-HF (i.e. non-selfconsistent orbitals).

# S.6 Frozen core approximation

The frozen core settings employed by default in ORCA, has the frozen core of 3d transition metals reduced by one shell than encountered more commonly. In this case, cobalt has a frozen core up to the 2p shell instead of 3p in the traditional frozen core approximation. The difference between the two, i.e. importance of correlating 3s and 3p electrons of cobalt, was tested at the cc-pVnZ-DK (N = 2-4) level with CPS cut-offs on the adiabatic ionization energy of cobaltocene. We also tested the effect of keeping 3s and 3p electrons frozen on the $\Delta E(T_1)$ and $\Delta E(CV)$ corrections. All $\Delta E(T_1)$ corrections are at the CBS(3/4) level with N.PNO cut-offs, and the $\Delta E(CV)$ correction was tested at both $T_0$ and $T_1$ levels at the QZ level with CPS cut-offs using the cc-pwCVnZ-DK for Co and cc-pVnZ-dK for C and H.

Keeping 18 electrons frozen (red) significantly increases the IE. We also see that for the smaller frozen core (blue), correlating core electrons at the $T_0$ (left-pointing triangle) or $T_1$ (right-pointing triangle) level makes no difference, unlike the larger frozen core. This indicates that $T_1$ triples are important for 3s and 3p electrons.

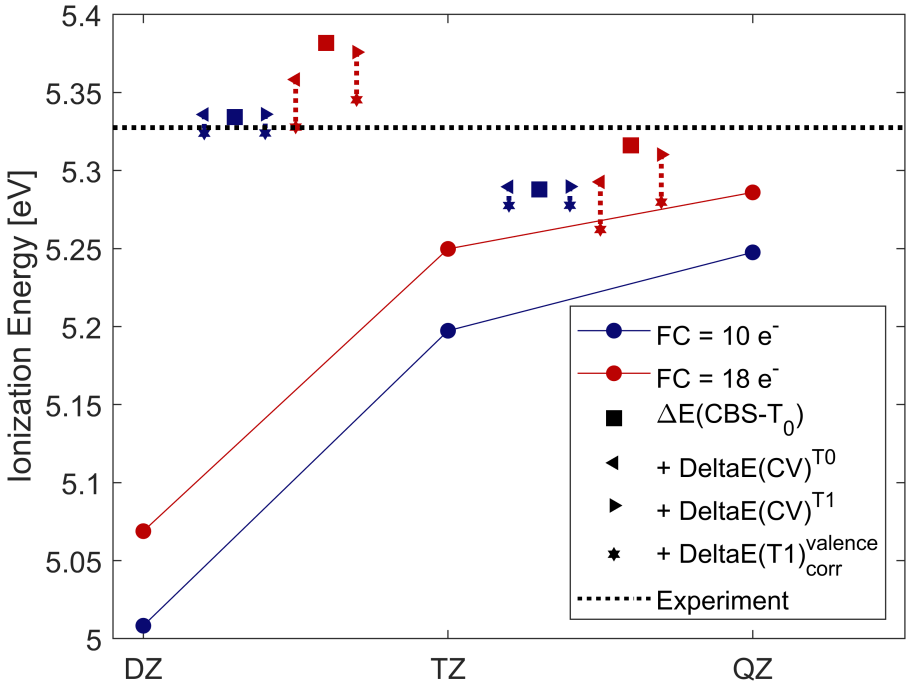


**Figure S.4** The line represents ionization energies with frozen cores of 18 $e^-$ (red) and 10 $e^-$ (blue), calculated with CPS cut-offs at the $T_0$ level with the cc-pVNZ-DK basis sets. The squares are CBS extrapolated ionization energies. Core-valence corrections, $\Delta E(CV)$, were added to CBS IEs at the $T_0$ (left-pointing triangle) and $T_1$ (right-pointing triangle) levels. The star shows the addition of the $\Delta E(T_1)$ correction for valence electrons.

Interestingly, the valence $\Delta E(T_1)$ correction (star) of active electrons for the larger frozen core is much larger, showing that the contribution of $T_1$ triples of 3s 3p electrons to the IE is in the opposite direction of $T_1$ for 3d electrons. This indicates that $T_1$ for both is needed for a balanced description.

This suggests that it should be valid to add a $\Delta E(CV)$ correction at the $T_0$ level and only adding a $\Delta E(T_1)$ correction to the active electrons when using the smaller frozen core. This strategy, on the other hand, leads to over-correction of the larger frozen core (red star below left triangle in CBS(3/4) IEs), indicating that using a larger frozen core would require a $\Delta E(CV)$ correction at the $T_1$ level.

Overall the results imply that the smaller frozen core definition for 3d transition metals (used by default in ORCA) makes CCSD(T) calculations overall more numerically robust and reduces the need for core-valence corrections, despite the larger associated cost.

# S.7 Calculations of vibrational frequencies and zero-point energy contributions

In the ZEKE MATI measurements, the spectrum had a well resolved vibronic structure for the $Cp_2Co^+$ ion, from which the ground state vibrational spectrum could be determined.[6] The authors highlighted four key vibrational modes, which were used in determining the IE. Harmonic vibrational calculations were carried out using the same DFAs as in the geometry optimizations with def2-TZVP basis sets, in order to evaluate how accurately they produce the four vibrational modes used in determining the experimental IEs. Table S.5 lists the calculated and experimental modes along with the zero-point energy contributions to the IE ($E(ZPE)_{Cp2Co^+} - E(ZPE)_{Cp2Co}$). The vibrational modes are labelled according to a symmetry convention used for metallocenes.[7]

**Table S.5** Experimental and calculated values for selected vibrational modes of $Cp_2Co^+$ and zero point energy contributions of the DFAs to the adiabatic ionizationenergy[a]

| | $\nu_2$ $a_1$' | $\nu_4$ $a_1$' | $\nu_{22}$ $e1$' | $\nu_{28}$ $e_2$' | $\Delta ZPE$[b] |
|---|---|---|---|---|---|
| **Expt.**[c] | **851** | **309** | **158** | **589** | |
| BP86 | 840 | 310 | 161 | 576 | 0.146 |
| TPSS | 861 | 315 | 162 | 576 | 0.150 |
| TPSSh | 876 | 318 | 165 | 601 | 0.148 |
| B3LYP | 870 | 301 | 160 | 602 | 0.140 |
| PBE0 | 879 | 317 | 167 | 615 | 0.141 |
| M06-2X | 880 | 290 | 160 | 631 | 0.108 |

[a] Vibrational frequencies are shown in $cm^{-1}$ zero-point energy contributions in eV. [b] $\Delta ZPE$ is the zero-point energy contribution to the adiabatic IE, $E(ZPE)_{Cp2Co^+}$ - $E(ZPE)_{Cp2Co}$. [c] From ref. 6

The experimentally best resolved mode is $\nu_4$, a totally symmetric stretching mode between Co and the Cp rings. Its progression through the quantum vibrational levels could be seen on the MATI spectrum, of which the vertical and adiabatic IPs could readily be identified. The $\nu_4$ peaks are always accompanied by three weaker peaks, namely $\nu_2$, $\nu_{22}$ and $\nu_{28}$.

The calculations are unfortunately complicated by the fact that anharmonicity is not accounted for in the calculations which should lower the harmonic frequencies. For $\nu_2$, an out of plane C-H bend, BP86 and TPSSh perform the best, with errors of -11 and 9 $cm^{-1}$, respectively. The error then increases with increasing fractional HF exchange. The out of plain tilting of the Cp rings, $\nu_{22}$, is well described by all methods, where B3LYP and M06-2X have the lowest errors, 2 $cm^{-1}$, and PBE0 has the largest error, 9 $cm^{-1}$. All methods apart from PBE0 and M06-2X have errors of similar magnitude for $\nu_{22}$, the out of plane ring-distortion responsible for the Jahn-Teller stabilization of the neutral $Cp_2Co$, although with opposite signs for hybrid DFAs. BP86 and TPSS both have errors of -13 $cm^{-1}$, whereas TPSSh and B3LYP have errors of 12 and 13 $cm^{-1}$, respectively, and PBE0 and M06-2X have somewhat larger errors of 26 and 42 $cm^{-1}$, respectively.

BP86 has the smallest error, 1 $cm^{-1}$, for the totally symmetric Co-Cp stretch, $\nu_4$. This is interesting when compared to PBE0, as these two DFAs predicted the best Co-C distance for the neutral $Cp_2Co$, but PBE0 predicted a greater contraction of the complex after ionization, resulting in a Co-C distance similar to that of TPSSh for the oxidized $Cp_2Co^+$ (table 2 - main text). This could be indicative of BP86 capturing the cation geometry more accurately, further supported by the fact, that DFAs which underestimated the Co-C distance of the neutral $Cp_2Co$, TPSS and TPSSh, also predict shorter Co-C distance for $Cp_2Co^+$, and generally seem to overestimate $\nu_4$ for $Cp_2Co^+$, and

vice versa.

All DFAs apart from M06-2X predict very similar ZPE contributions to the IP (within 10 meV). M06-2X predicts a much smaller contribution, likely related to the large differences in optimized geometries. Another detail to consider are low frequency large amplitude motions (LAMs), commonly observed in non-rigid molecules. The vibrational mode with the lowest frequency for both species is the rotation of the Cp rings, predicted below 50 $cm^{-1}$ by all DFAs. Additionally, all DFAs tested here, except M06-2X, predict one of the $\nu_{28}$ modes of the neutral $Cp_2Co$ at much lower frequency, roughly 50 $cm^{-1}$, likely due to the Jahn-Teller distortion. These LAMs are generally anharmonic and likely poorly described by the QRRHO approximation. This will subsequently affect the vibrational entropy contributions, not shown here, but will also have some effect on the ZPE. LAMs are notoriously difficult to treat accurately, and outside of the scope of this study.[8]

# S.8 Relativistic effects

The scalar relativistic DKH Hamiltonian and the appropriate relativistic basis sets were included throughout all calculations in the main manuscript. Here, however, we analyze directly the effect of scalar relativity on the adiabatic IE of $Cp_2Co$ by comparing non-relativistic, ZORA and DKH scalar relativistic Hamiltonians in DLPNO-CCSD($T_0$) CBS(3/4) calculations with N.PNO cut-offs and BP86 reference orbitals, in addition to several density functionals. The effect of the ZORA or DKH Hamiltonian was calculated by taking the difference of a calculation with a relativistic Hamiltonian and a relativistically recontracted basis set on one hand and a nonrelativistic Hamiltonian with the non-relativistic basis set version on the other. Results are shown in Table S.6.

**Table S.6** Ionization energies calculated with/without ZORA and DKH scalar relativistic effects[a]

| | Non-Rel. | DKH | $\Delta_{DKH}$ | ZORA | $\Delta_{ZORA}$ |
|---|---|---|---|---|---|
| CCSD($T_0$)(def2)[b] | 5.3775 | 5.3242 | -0.0533 | 5.3169 | -0.0605 |
| CCSD($T_0$)(cc)[c] | 5.3859 | 5.3024 | -0.0834 | - | - |
| BP86 | 5.3227 | 5.2570 | -0.0657 | 5.2501 | -0.0726 |
| TPSS | 5.1188 | 5.0501 | -0.0686 | 5.0432 | -0.0755 |
| TPSSh | 5.2618 | 5.1889 | -0.0729 | 5.1820 | -0.0798 |
| B3LYP | 5.5110 | 5.4372 | -0.0738 | 5.4302 | -0.0809 |
| PBE0 | 5.5820 | 5.5033 | -0.0787 | 5.4962 | -0.0858 |

[a] All energies are shown in eV and include a 0.146 eV ZPE contribution. The DFT calculations use the nonrelativistic def2-TZVP basis set or the relativistically recontracted DKH or ZORA version.[b] DLPNO-CCSD($T_0$), CBS(3/4), NormalPNO, with (ZORA/DKH-)def2-(T/Q)ZVPP basis sets. [c] DLPNO-CCSD($T_0$), CBS(3/4), NormalPNO, with cc-pV(T/Q)Z(-DK) basis sets.

The results suggest that there is a sizeable relativistic effect of -0.05 to -0.08 eV depending on the relativistic Hamiltonian and which basis set family is used. The relativistic effect is similar for both CCSD(T) and DFT methods. In order to confirm that the effect on the IE is a direct effect of relativity, rather than a results of the different basis set contraction of the core-region, we also performed BP86 calculations with flexible fully decontracted def2-SVP, def2-TZVPP and def2-QZVPP basis sets. The comparison between non-relativistic calculations and ZORA/DKH with decontracted basis sets can be seen in Table S.7

The relativistic effect on the IE is the same when calculated with identical decontracted def2 basis sets for both relativistic and non-relativistic calculations as when specially recontracted DKH/ZORA-def2 basis sets are used, as confirmed by comparison of Tables S.6 and S.7. This confirms that the effect is indeed scalar relativistic in nature, and not due to the different basis set

**Table S.7** Ionization energies calculated using identical decontracted def2 basis sets for relativistic and non-relativistic BP86 calculations[a]

| | SVP | $\Delta_{Rel}$ | TZVPP | $\Delta_{Rel}$ | QZVPP | $\Delta_{Rel}$ |
|---|---|---|---|---|---|---|
| Non-Rel. | 5.2474 | | 5.3545 | | 5.3593 | |
| DKH | 5.1814 | -0.0660 | 5.2881 | -0.0664 | 5.2928 | -0.0665 |
| ZORA | 5.1745 | -0.0729 | 5.2811 | -0.0734 | 5.2858 | -0.0735 |

[a] All energies are shown in eV and include a 0.146 eV ZPE contribution.

**Table S.8** Individual DLPNO-CCSD($T_0$) (CBS(3/4), N.PNO) contributions to the ionization energy of $Cp_2Co$ with and without scalar relativistic effects for both def2-(T/Q)ZVP and cc-pV(T/Q)Z basis sets[a]

| | Ref. | CCSD | ($T_0$) | Total | $\Delta_{Rel.}$[b] |
|---|---|---|---|---|---|
| cc-non-rel. | 5.3966 | 0.0920 | -0.2495 | 5.2392 | |
| cc-DKH | 5.2975 | 0.1022 | -0.2440 | 5.1557 | -0.0834 |
| def2-non-rel. | 5.3981 | 0.0643 | -0.2316 | 5.2308 | |
| def2-DKH | 5.2927 | 0.1161 | -0.2314 | 5.1775 | -0.0533 |
| def2-ZORA | 5.2853 | 0.1161 | -0.2312 | 5.1702 | -0.0605 |

[a] All energies are shown in eV. Calculations use BP86 reference orbitals. No ZPE contributions were included.[b] Energy relative to the non-relativistic calculations from the top row for each basis set.

contraction used in non-relativistic vs. relativistically recontracted basis sets.

For the DLPNO-CCSD(T) calculations, the reference, CCSD and ($T_0$) contributions were compared in relativistic vs. non-relativistic calculations. It is evident from the comparison of relativistic contributions to the IE in Table S.8, that the relativistic effect mainly stems from the SCF energy (ca. -0.1 eV), although there is a slight contribution in the opposite direction from CCSD correlation.

Valence MO diagrams for QROs of neutral $Cp_2Co$ calculated with BP86 using both cc and def2 basis sets with ZORA and DKH are compared in Figure S.5. Hardly any difference is observed between valence MOs calculated with or without scalar relativistic Hamiltonian. The reason for the surprisingly large relativistic effect on the IE of a 3d metal complex is therefore not obvious and likely stems from a general effect in the core region.

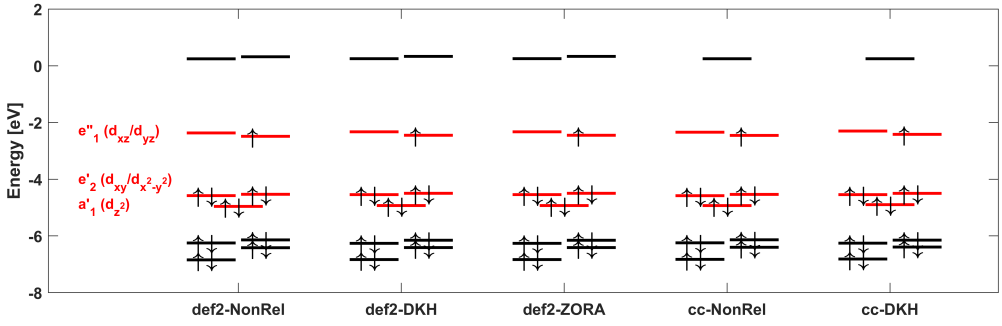


**Figure S.5** QRO energies of $Cp_2Co$ calculated with and without scalar relativistic effects. All calculations were done with BP86. Both def2-TZVP and cc-pVTZ were used along with their relativistic counterparts.

# S.9 PNO analysis for different reference determinants

Recent spin-state energetics studies by Pantazis and coworkers on Co and Mn complexes calculated with DLPNO-CCSD(T), found that the use of BP86 orbitals consistently gave the best results.[9,10] In these studies, mean- and maximum- PNOs included per electron pair were compared between the different reference determinants. References leading to more PNOs per pair being included in the calculation generally gave more accurate results. This was interpreted as being related to different orbital shapes, which affects the way domains are defined in the DLPNO treatment.

Table S.9 shows the mean- and maximum- PNOs per pair in cc-pVQZ-DK calculations of cobaltocene with different reference orbitals along with their respective DLPNO-CCSD($T_0$) errors in IE from CBS(3/4) calculations (from Fig. 3 - main text). There is generally a good correlation between the errors in calculated IE and the mean/max PNOs per pair. The meta-DFAs deviate slightly from the observed correlation, perhaps due to the higher electron density of cobalt (Fig. 5 - main text) and subsequent lowering in IE for TPSS and TPSSh. The mean/max PNOs per pair are very close for all DFT references in calculations of $Cp_2Co^+$, whereas they differ more for the neutral. Consistent with the studies by Pantazis et al., BP86 reference orbitals (and other GGAs) include the the highest number of PNOs in the DLPNO-CCSD calculations and lead also to the lowest errors. The mean/max PNOs per pair then seems to lower systematically with increasing HF exchange in the DFA, as does the accuracy. The HF reference leads to the lowest number of PNOs in correlated calculations for both the neutral and cation, and has much larger errors in adiabatic IE than the rest. Based on these results, analysis of mean/max PNOs per pair using smaller basis sets may be useful for selecting suitable reference orbitals for DLPNO-CC calculations of a given system.

**Table S.9** Average- and maximum- number of PNOs per pair in cc-pVQZ-DK DLPNO-CCSD calculations of $Cp_2Co$ and $Cp_2Co^+$ using different reference orbitals compared to errors in IE from DLPNO-CCSD($T_0$) calculations.

| | **$Cp_2Co$** | | **$Cp_2Co^+$** | | |
|---|---|---|---|---|---|
| | Avg. PNO | Max. PNO | Avg. PNO | Max. PNO | $\Delta$IE[a] |
| BP86 | 41 | 117 | 41 | 104 | −26 |
| TPSS | 41 | 116 | 41 | 104 | −36 |
| TPSSh | 40 | 108 | 41 | 104 | −43 |
| B3LYP | 40 | 106 | 40 | 104 | −37 |
| PBE0 | 40 | 104 | 40 | 104 | −44 |
| M06-2X | 38 | 104 | 40 | 104 | −51 |
| BHLYP | 38 | 104 | 39 | 104 | −67 |
| HF | 37 | 103 | 37 | 102 | −137 |

[a] Error in IE relative to experiment from the DLPNO-CCSD($T_0$) calculations in Fig 3 (main text), shown in meV.

# S.10 Computational cost

Unfortunately, we can only report rough estimates on computational cost, as calculations had to be run on different nodes that do not have the same amount of RAM, depending on availability at the cluster. The most demanding calculation by far for the $\Delta E(T_1)$ corrections, is the QZ-T.PNO calculation (same as tighter cut-off in CPS) of open-shell $Cp_2Co$, which took 7.5 days using 6 cores

and 32 GB RAM per core, or roughly 6 times longer than the closed-shell $Cp_2Co^+$ on equivalent nodes. Using a node with 50% less RAM available, the entire N.PNO CBS(3/4) $\Delta E(T_1)$ correction (4 calculations) can be calculated in roughly the same amount of time as only the QZ-T.PNO calculation of neutral $Cp_2Co$, which takes about twice as long as the L.PNO calculation on that same same node. Furthermore, at the TZ-CPS level, the correction is almost 50% more expensive than N.PNO CBS(3/4). For comparison, the all electron calculation of neutral open-shell $Cp_2Co$ in $\Delta E(CV)_{corr}$ with cc-pwCVNZ basis sets on all atoms (non-relativistic) took 5 days using 6 cores and 32 GB RAM per core. Using the cc-CV/V basis sets instead significantly reduces the computational cost. Fortunately, $\Delta E(T_1)$ does not appear to be required for correlation of core electrons as all-electron calculations at the $(T_1)$ level could be prohibitively expensive, especially for larger systems. Furthermore, the $(T_1)$ correction for valence electrons can also to a good approximation be described with a smaller basis set and PNO cutoff as shown below.

# S.11 Study of the basis set and PNO convergence of the $T_1$ correction for ferrocene

The dependence of the $T_1$ correction on basis set size and PNO cutoffs was studied separately for ferrocene as it is an order of magnitude larger compared to the $T_1$ correction for cobaltocene. The results are shown in S.10. The results reveal that the correction is not overly dependent on basis set or PNO cutoffs and can to a good approximation be calculated with e.g. a TZ basis set and NormalPNO setting but with a reduced $T_{CutPNO}$ ($10^{-6}$) cutoff.

**Table S.10** Convergence of the $\Delta E(T_1)_{corr}$ term for $Cp_2Fe$ with respect to PNO cutoffs and basis set size using cc-pwCVnZ for Co, cc-pVnZ for C and H and PBE0 reference orbitals[a]

| | DZ | TZ | QZ | DZ/TZ | TZ/QZ |
|---|---|---|---|---|---|
| LoosePNO | 0.188 | 0.191 | 0.205 | 0.193 | 0.215 |
| NormalPNO*[b] | 0.218 | 0.222 | 0.227 | 0.224 | 0.231 |
| NormalPNO | 0.225 | 0.234 | 0.242 | 0.240 | 0.248 |
| TightPNO*[b] | 0.218 | 0.222 | 0.227 | 0.224 | 0.231 |
| TightPNO | 0.230 | 0.243 | 0.247 | 0.251 | 0.250 |

[a] All energies are shown in eV.[b] PNO cut-offs have a reduced $T_{CutPNO}$ ($1x10^{-6}$), with other cut-offs unchanged.

# S.12 ASH scripts

The pre-programmed CCSD(T) workflow feature in the Python-based ASH program, via a flexible ORCA interface, allow one to conveniently perform the final protocol discussed in the article. A Python object of class "ORCA_CC_CBS_Theory" with various input options is created and single-point calculations can then be performed for a single molecule or multiple molecules in a single script.

See also documentation at: ASH highlevel workflows and ASH highlevel workflows tutorial

Example ASH script for running a DLPNO-CCSD(T)/CBS calculation on cobaltocene is shown below. This example performs a CBS(3,4) extrapolation of the correlation energy using the cc-pVnZ-DK basis set family at the DLPNO-CCSD($T_0$) level of theory, without a core-valence correction (CVSR=False), with a $(T_1)$-correction (T1correction=True) using the smaller basis set for $T_1$ correction (T1corrbasis_size='Small' meaning cc-pVTZ-DK here), using reduced NormalPNO setting

(TCutPNO=1e-6), PNO extrapolation according to the CPS extrapolation formula by Drosou et al. A DKH Hamiltonian is used (relativity='DKH'), with a KS-DFT(BP86) orbital reference (DFTreference="BP86")and the SCF reference energies are not extrapolated (SCFextrapolation=False).

```
from ash import *

numcores=24 #Number of cores reserved
actualcores=16 #Number of cores used

#Defining molecular fragments
cpco0=Fragment(xyzfile="CpCo_0_gas.xyz", charge=0, mult=2)
cpcoI=Fragment(xyzfile="CpCo_I_gas.xyz", charge=1, mult=1)
# Defining species, stoichiometry and reaction
specieslist=[cpco0,cpcoI]
stoichiometry=[-1, 1]
reaction = Reaction(fragments=specieslist, stoichiometry=stoichiometry)

#Defining a ORCA_CC_CBS_Theory object
cc = ORCA_CC_CBS_Theory(elements=cpco0.elems, cardinals=[3,4], basisfamily="cc-dk",
    DFTreference="BP86", DLPNO=True, CVSR=False, T1correction=True, T1corrbasis_size='
    Small', T1corrpnosetting='NormalPNOreduced', numcores=actualcores, pnosetting="
    extrapolation", pnoextrapolation=[1e-6,3.33e-7,2.38,'NormalPNO'], memory=20000,
    scfsetting="Verytightscf", relativity='DKH', SCFextrapolation=False)

Singlepoint_reaction(theory=cc, reaction=reaction, unit='eV')
```

Shown below is also an example ASH script for running the whole metallocene benchmark set using the run_benchmark feature of ASH:

```
from ash import *

numcores=24
elements=['C','H','Cr','V','Co','Mn','Ni','Fe']
#Define ORCA_CC_CBS_Theory object
cc = ORCA_CC_CBS_Theory(elements=elements, cardinals=[2,3,4], basisfamily="cc-dk",
    DFTreference="BP86", DLPNO=True, Triplesextrapolation=True, CVSR=False, T1correction=
    True, T1corrbasis_size='Small', T1corrpnosetting='NormalPNOreduced', numcores=numcores
    , pnosetting="extrapolation", pnoextrapolation=[1e-6,3.33e-7,2.38,'NormalPNO'], memory
    =21000, scfsetting="Verytightscf", SCFextrapolation=False, relativity='DKH')

run_benchmark(set="Metallocenes", theory=cc, numcores=12)
```

# Bibliography


[1] A. Karton and J. M. Martin, *Theoretical Chemistry Accounts*, 2006, **115**, 330–333.

[2] F. Neese and E. F. Valeev, *Journal of chemical theory and computation*, 2011, **7**, 33–43.

[3] T. Helgaker, W. Klopper, H. Koch and J. Noga, *The Journal of chemical physics*, 1997, **106**, 9639–9646.

[4] B. M. Flöser, Y. Guo, C. Riplinger, F. Tuczek and F. Neese, *Journal of chemical theory and computation*, 2020, **16**, 2224–2235.

[5] T. D. Crawford and H. F. Schaefer, *The Journal of Chemical Physics*, 1996, **104**, 6259–6264.

[6] S. Y. Ketkov and H. L. Selzle, *Angewandte Chemie International Edition*, 2012, **51**, 11527–11530.

[7] J. Bodenheimer and W. Low, *Spectrochimica Acta Part A: Molecular Spectroscopy*, 1973, **29**, 1733–1743.

[8] C. Puzzarini, J. Bloino, N. Tasinato and V. Barone, *Chemical reviews*, 2019, **119**, 8131–8191.

[9] S. E. Neale, D. A. Pantazis and S. A. Macgregor, *Dalton Transactions*, 2020, **49**, 6478–6487.

[10] M. Drosou, C. A. Mitsopoulou and D. A. Pantazis, *Polyhedron*, 2021, **208**, 115399.