ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/44677989

# Spin-State-Corrected Gaussian-Type Orbital Basis Sets

**ARTICLE** *in* THE JOURNAL OF PHYSICAL CHEMISTRY A · JULY 2010

Impact Factor: 2.69 · DOI: 10.1021/jp102712z · Source: PubMed

CITATIONS
22

READS
28

**4 AUTHORS**, INCLUDING:

Marcel Swart
Catalan Institution for Research and Advan…

**155** PUBLICATIONS **3,431** CITATIONS

SEE PROFILE

Josep M. Luis
Universitat de Girona

**106** PUBLICATIONS **2,456** CITATIONS

SEE PROFILE

Miquel Solà
Universitat de Girona

**398** PUBLICATIONS **8,667** CITATIONS

SEE PROFILE

Available from: Josep M. Luis
Retrieved on: 25 January 2016

# Spin-State-Corrected Gaussian-Type Orbital Basis Sets

**Marcel Swart,*,†,‡ Mireia Güell,† Josep M. Luis,† and Miquel Solà†**

*Institut de Química Computacional and Departament de Química, Universitat de Girona, Campus Montilivi, 17071 Girona, Spain, and Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, 08010 Barcelona, Spain*

*Received: March 25, 2010; Revised Manuscript Received: May 17, 2010*

Recently, we reported that the basis set has a profound influence on the computed values for spin-state splittings [*J. Phys. Chem. A* **2008**, *112*, 6384]. In particular, small Gaussian-type orbital (GTO) basis sets were shown to be unreliable for the prediction of them. Here, we report simple modifications of the small Pople-type Gaussian-type orbital basis sets (3-21G, 3-21G*, 6-31G, 6-31G*), which correct their faulty behavior for the spin-state energies. We have investigated the basis sets for a set of 13 first-row transition-metal complexes for which reliable reference data have been obtained at the OPBE/TZ2P(STO) level. For several systems, we have used single and double spin-contamination corrections to avoid ambiguity of the results because of spin contamination, that is, the energies and geometries were obtained for the pure spin states. The spin ground states as predicted by the spin-state-corrected GTO basis sets (s6-31G, s6-31G*) are in complete agreement with the reference Slater-type orbital (STO) data, while those of the original basis sets and a recent modification by Baker and Pulay (m6-31G*) are not for all cases. The spin-state-corrected GTO basis sets also improve upon the original and modified basis sets for the accuracy of geometry optimization, while the accuracy of the vibrational frequencies is as good or better. At a limited additional cost, one therefore obtains very reliable results for these important spin-state energies.

## Introduction

Spin-state energies of transition-metal complexes form a vital part of different stages in the catalytic cycle of metalloenzymes.[1] The best example for this is posed by the cytochrome P450cam, in which the spin state goes from low-spin doublet to high-spin sextet, high-spin quintet, low-spin singlet, and may in the rate-determining steps have either a low-spin doublet or a quartet ground state.[2,3] To understand the mechanism for the catalytic cycle, and how it is affected by changes in the substrate or active-site residues, theory plays an important role.[3–5] However, it was shown that the correct prediction of spin-state splittings is a difficult task for both density functional and wave function-based methods.[1,4–6]

Apart from the choice of methodology, also the choice of basis set used for the calculations was shown to be vitally important.[1,7] Good behavior was observed for Slater-type orbitals (STOs) and large Gaussian-type orbitals (GTOs), which converged to the same result. The rate of convergence was, however, much faster for the STOs than for the GTOs. Small and medium-sized GTOs in particular performed very poorly especially for high-spin states that are usually predicted at too low energy. As a result, these latter GTOs often predict the wrong spin ground state as is the case, for example, for FeFHOH, one of the compounds used in our previous study.[7] Previously, we attributed this failure of, for example, 3-21G in part to the limited number of GTO primitives for the core region,[7] which is a notion we will investigate here in more detail.

It should be stressed that basis sets involving effective core potentials (ECPs), which together with the corresponding valence basis sets are referred to as ECP basis sets (ECPBs), give a systematic and worrying disagreement with the reliable STO/large-GTO data.[7] Therefore, the use of these ECPBs for spin-state energies is not recommended.[1,7]

To find out which is the origin of the problems with the small and medium-sized GTO basis sets, we have investigated in detail the performance for spin-state splittings and propose here small modifications that largely help to solve their problems. The resulting spin-state-corrected GTO basis sets have been used for a set of transition-metal complexes for which reference data have been obtained with a large TZ2P basis set of Slater-type functions. This latter basis set was previously[7] shown to give reliable results. Here, we investigate not only the performance for spin-state energies but also that for geometries and vibrational frequencies.

## Computational Details

All spin-state energies for the transition-metal complexes studied have been obtained with the OPBE[8–10] density functional approximation, which was recently[5] shown to be the most reliable density functional (together with the recently developed SSB-D[6] functional) for spin-state splittings. The spin-state splittings with the GTO basis sets have been obtained with the NWChem program[11] (version 5.1.1) using the spin-state-relaxed geometries that have been obtained at the OPBE/TZ2P(fc) level using the ADF[12] (2009.01) and QUILD[13] (2009.01a) programs. The ADF program has been used only for providing the energy and gradient, while the actual geometry optimizations have been carried out with the QUILD program (included within the ADF program package) that has an improved geometry optimization scheme using adapted delocalized coordinates.[14] The computed spin-state splittings correspond to the relaxed (adiabatic) situation in which each spin state is obtained in its own "relaxed" optimized geometry.

* To whom correspondence should be addressed.
† Universitat de Girona.
‡ Institució Catalana de Recerca i Estudis Avançats.

10.1021/jp102712z © 2010 American Chemical Society
Published on Web 06/16/2010

Several compounds showed severe spin contamination as indicated by the expectation value for $S^2$, which should be 0, 2, and 6 for pure singlet, triplet, and quintet states and 0.75, 3.75, and 8.75 for pure doublet, quartet, and sextet states. The contaminated energy can be corrected by taking into account the next state with higher multiplicity.

$$E_{\text{pure}} = \frac{E_{\text{cont}} - a_1 \cdot E_{S+1}}{1 - a_1}$$
$$a_1 = \frac{\langle S^2 \rangle_{\text{cont}} - S \cdot (S+1)}{\langle S^2 \rangle_{S+1} - S \cdot (S+1)} \quad (1)$$

However, this equation is only valid for the case that this next state is actually a pure state. For several of the cases studied, which in all cases happened with the GTO basis sets, it was observed that these next states were also significantly spin-contaminated (e.g., with an $S^2$-value of 2.65 instead of 2.0). Therefore, these next-higher states should also be corrected resulting in a double spin-contamination correction.

$$a_2 = \frac{\langle S^2 \rangle_{S+1} - (S+1) \cdot (S+2)}{\langle S^2 \rangle_{S+2} - (S+1) \cdot (S+2)}$$
$$E_{\text{corr},S+1} = \frac{E_{\text{cont},S+1} - a_2 \cdot E_{S+2}}{1 - a_2}$$
$$a_3 = \frac{\langle S^2 \rangle_{\text{cont}} - S \cdot (S+1)}{(S+1) \cdot (S+2) - S \cdot (S+1)}$$
$$E_{\text{pure}} = \frac{E_{\text{cont}} - a_3 \cdot E_{\text{corr},S+1}}{1 - a_3} \quad (2)$$

The first of these corrections (eq 1) can be applied within the QUILD program[13] directly also for geometry optimizations.

## Results and Discussion

To get a deeper understanding of the origin of the deficiencies of the small GTO basis sets, we started by looking at the spin-state splitting for a small molecule, FeFHOH, that was used already in our previous study.[7] For this molecule, we investigated how the spin-state splittings with the 6-31G* basis set[15] changed by adding s-, p-, and d-functions for iron. Because the spin states mainly involve changes in the occupation of 3d-orbitals, it was to be expected that the d-functions would be responsible for the differences in spin-state energies. Indeed, we found that adding s- or p-functions (either diffuse or tight) does not lead to any change in the spin-state splittings. Of course, the addition of these functions does have a significant influence on the total energy of the compound, but the effect is the same for all three spin states considered. Similarly, the addition of a tight d-function to iron only has a small effect (less than 0.1 kcal·mol$^{-1}$) on the spin-state splittings.

A totally different pattern was observed when we added a diffuse d-function to the 6-31G* basis set. For instance, by adding a d-function with exponent 0.14275, the quartet−sextet energy difference for FeFHOH changed from −0.7 kcal·mol$^{-1}$ to ca. +10.4 kcal·mol$^{-1}$. This latter value is very close to the reference STO/large-GTO data of ca. 9.0 kcal·mol$^{-1}$, that is, only 1 kcal·mol$^{-1}$ higher in energy. Moreover, it now predicts directly the correct spin ground state. Adding a second additional d-function (with exponent 0.04041) did not have any significant effect (less than 0.1 kcal·mol$^{-1}$ difference for the sextet−quartet splitting). The situation for this second additional d-function may be similar to what we observed before with the cc-pVTZ and cc-pVQZ basis sets that gave similar results.[7]

**TABLE 1: Exponent Values[a] for Additional d-Functions in the s3-21G, s3-21G*,[b] s6-31G, and s6-31G* Basis Sets**

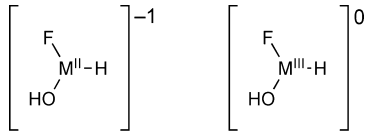
| | s3-21G, s3-21G* | s6-31G, s6-31G* |
|---|---|---|
| Sc | 0.07662 | 0.06020 |
| Ti | 0.10013 | 0.07931 |
| V | 0.12083 | 0.09606 |
| Cr | 0.14131 | 0.11248 |
| Mn | 0.16429 | 0.13050 |
| Fe | 0.18060 | 0.14275 |
| Co | 0.19985 | 0.15765 |
| Ni | 0.22056 | 0.17393 |
| Cu | 0.24226 | 0.19021 |
| Zn | 0.26569 | 0.20851 |

[a] In all cases, the exponents are added with weight of 1.0 (see Supporting Information for an example). [b] The s3-21G* contains a polarization f-function with exponent 0.8 similar to 6-31G*.

The exponent of 0.14275 (and 0.04041) that we used in this example was obtained in an even-tempered fashion and is based on the three smallest exponents of the d-function primitives for iron in 6-31G* (6.122368, 1.846601, 0.504361). For this series of three exponents, the exponent values decrease on average by a factor of 0.283 in each of the two steps. By now multiplying the lowest exponent with the same factor, we obtain the next even-tempered exponent of 0.14275, which was added with a weight of 1.0. In similar fashion, we have obtained exponent values for additional d-functions for the first-row transition-metal atoms Sc−Zn for both the s3-21G/s3-21G* and s6-31G/s6-31G* basis sets (see Table 1 and Supporting Information).

The notion that the representation of the 3d orbitals for transition-metal atoms has to be augmented has been described earlier, for example, by Hay in 1977[16] and more recently by Mitin et al.[17,18] Hay investigated (4d) and (5d) GTO expansions of the 3d orbitals of transition metals and contraction schemes for these expansions. It was observed[16] that a [3d] contraction scheme gave consistent results and was significantly more accurate than a [2d] contraction scheme. The study by Baker, Pulay and co-workers[17,18] modified the 6-31G* basis set (to give m6-31G*) and also observed that one of the main problems of 6-31G* is the lack of a sufficiently diffuse d-function on the metal. Similar to our attempts mentioned above, Baker and Pulay added a diffuse d-function.[17,18] However, they then reoptimized the complete set of exponents for each of the transition metals thereby keeping the [2d] contraction scheme of the 6-31G* basis intact. Although this is strictly speaking correct, it means a deterioration of the results in practice (vide infra) as already advocated previously by Hay.[16] In a previous study,[19] the optimization of the polarization (d) exponents of the 6-31G* basis set in reactants, transition states, and products of a couple of reactions led to optimized polarization functions that were also somewhat more diffuse than the standard values.

Here, we follow a different strategy in which we simply add a diffuse d-function in an even-tempered fashion, that is, giving the preferred [3d] contraction.[16] Strictly speaking, this means that we no longer have a double-$\zeta$ representation for the d-functions but triple-$\zeta$. However, the s- and p-functions are still the same, so we keep on referring to the spin-state-corrected basis sets as double-$\zeta$ basis sets, for example, s6-31G*. Later on, we will return to this issue where it will be shown that having this [3d] contraction scheme for the metal 3d orbitals is actually needed.




MFHOH⁻ MFHOH

**Figure 1.** Small transition-metal complexes studied here.

**TABLE 2: Reference Spin-State Splittings (OPBE/TZ2P, kcal·mol$^{-1}$) for M$^{II}$FHOH$^-$ and M$^{III}$FHOH**

| | M$^{II}$FHOH$^-$ | | | M$^{III}$FHOH | | |
|---|---|---|---|---|---|---|
| | low | interm. | high | low | interm. | high |
| Ti | 9.5 | 0[a] | | | | |
| V | 22.9[b] | 0[c] | | 29.5 | 0[a] | |
| Cr | 67.7[d] | 35.9[b] | 0[e] | 39.4[b] | 0[c] | |
| Mn | 70.9[b] | 26.3 | 0[f] | 77.7[d] | 39.3 | 0[e] |
| Fe | 48.2[d] | 19.1 | 0[e] | 38.1[b] | 0[c] | 8.8 |
| Co | 4.1 | 0[c] | 61.2 | 20.0[d] | 0[a] | 17.0 |
| Ni | 0[g] | 4.5 | | 0[h] | 31.2 | |

[a] Spin ground state, triplet. [b] Spin contamination corrected energy. [c] Spin ground state, quartet. [d] Obtained by using augmented Roothaan−Hall (ARH) SCF, see text. [e] Spin ground state, quintet. [f] Spin ground state, sextet. [g] Spin ground state, singlet. [h] Spin ground state, doublet.

**TABLE 3: Mean Absolute Deviation (MAD) and Maximum Deviation (MAX) for Different GTO Basis Sets with Respect to Reference STO-TZ2P Data[a]**

| basis set | MAD | MAX | # wrong g.s.[b] |
|---|---|---|---|
| 3-21G | 7.09 | 23.03 | 3 |
| 3-21G*[c] | 7.80 | 24.92 | 4 |
| 6-31G | 4.11 | 13.73 | 2 |
| 6-31G* | 3.60 | 12.67 | 2 |
| m6-31G | 1.98 | 12.40 | 1 |
| m6-31G* | 1.61 | 6.31 | 1 |
| **basis set proposed here** | **MAD** | **MAX** | **# wrong g.s.[b]** |
| s3-21G | 3.46 | 15.83 | 1 |
| s3-21G* | 3.61 | 16.89 | 1 |
| s6-31G | 1.47 | 10.01 | 0 |
| s6-31G* | 1.06 | 5.20 | 0 |
| cs3-21G*[d] | 3.41 | 16.94 | 1 |
| cs6-31G*[d] | 1.36 | 7.53 | 0 |

[a] All in kcal·mol$^{-1}$. [b] Number of wrong ground states predicted by the basis set for the transition-metal complexes. [c] The 3-21G* basis set contains a polarization f-function with exponent 0.8 similar to 6-31G*. [d] Diffuse d-function (see Table 1) mixed in with smallest standard d-function with weight 0.2 and 0.8 for diffuse and smallest function, respectively.

**Relaxed Spin-State Splittings for M$^{II}$FHOH$^-$ and M$^{III}$FHOH.** In our previous study on the influence of the basis set on spin-state splittings, we used a small compound, FeFHOH, for which the self-consistent field (SCF) converges easily. Moreover, because it has no or low symmetry within the different spin states, no complications arise from having spin states with different orbital occupations of different irreps, which might be close in energy. This latter is, for instance, the case for $FeF_3$ that has several quartet states that are only 3−11 kcal·mol$^{-1}$ apart and that are difficult to obtain straightforwardly with quantum-chemistry programs that do not have the option of explicitly setting the occupation of the several irreps.

Because of the small size of this FeFHOH compound, we explored here the relaxed (adiabatic) spin-state splittings of a set of similar complexes (see Figure 1) with several first-row transition-metal atoms (Ti−Ni) in different redox states (II, III) and for which different spin states are possible. The reference data for the spin-state splittings have been obtained using the OPBE functional with the all-electron STO-TZ2P basis set (see Table 2), which was previously[1,4,5,7] shown to give reliable and consistent spin-state energies. In this study, the choice of the density functional is not important as we are concerned here only with the basis set. Nevertheless, the OPBE functional[8] and its successor (SSB-D)[6] are by now well-known[1,5,20] to give reliable spin-state splittings. Also, for four systems (singlet states of Co$^{III}$, Cr$^{II}$, Fe$^{II}$, Mn$^{III}$), we used the augmented Roothaan−Hall (ARH) equations[21] as implemented in ADF, which solved the problems of difficult SCF for these compounds. Furthermore, spin-contamination corrections (eq 1) were carried out for several systems both for the energy and for the geometry (see Table 2).

With the STO-TZ2P optimized geometries, we performed single-point energy calculations with a range of (small) Pople-type GTO basis sets. The deviations of these energies from the reference STO-TZ2P energies are summarized in Table 3, while Table 4 shows the energies for all basis sets and compounds. Table 3 shows clearly the problems of the original small GTO basis sets, which have a mean absolute deviation (MAD) between 3.6 and 7.1 kcal·mol and a maximum absolute deviation (MAX) between 12.7 and 24.9 kcal·mol$^{-1}$ for 3-21G, 3-21G*, 6-31G, and 6-31G*. Moreover, the wrong spin ground state is predicted for at least two and a maximum of four complexes. Because the ARH approach is not yet available within the NWChem code that we use for the results with the GTO basis sets, we do not take the four compounds with problematic SCF (see above) into account when comparing the deviation of the GTO basis sets. Several of these compounds needed a double spin contamination (see eq 2). The poor performance of the original GTO basis sets is indeed drastically improved by the modifications of Mitin et al. (m6-31G, m6-31G*).[18] The MAD value drops to below 2.0 kcal·mol$^{-1}$ (see Table 3), and the maximum deviation also reduces significantly especially for m6-31G*. However, these modified basis sets still are unable to predict the correct spin ground state for all transition-metal complexes. In particular, they are unable to predict the singlet ground state of Ni$^{II}$FHOH (see Table 3).

The basis sets proposed here (e.g., s6-31G*) do an even better job than the modified ones (e.g., m6-31G*). The improvement of the spin-state-corrected s3-21G and s3-21G* basis sets over the corresponding original basis sets (e.g., 3-21G) is considerable with a reduction of the MAD values by approximately half. The maximum deviation also reduces considerably, and for these small basis sets, only one wrong spin ground state is predicted (that of $^2$Co$^{II}$FHOH$^-$). Overall, these corrected basis sets are thus approximately as good as 6-31G* (see Table 3). Our previous assessment that the problem of the 3-21G/3-21G* basis set should be coming from its insufficient description of the core region must thus be revised. Here, we show that it is "merely" the 3d orbital space that is at fault.

Surprisingly good results are obtained with the s6-31G and s6-31G* basis sets. The MAD and maximum deviation drop even further, that is, to 1.1 kcal·mol$^{-1}$ (MAD) and 5.2 kcal·mol$^{-1}$ (MAX) for s6-31G*. Moreover, both basis sets correctly predict the spin ground state for all transition-metal complexes. On the basis of this good performance and its still small size, we recommend the use of the s6-31G* basis set as a reliable and fast basis set for determining spin-state energies within GTO-based quantum-chemistry programs.

**TABLE 4: Spin-State Splittings (kcal·mol$^{-1}$) As Obtained with a Range of GTO Basis Sets**

| M | *TZ2P*[a] | 3-21G | 3-21G* | 6-31G | 6-31G* | m6-31G | m6-31G* |
|---|---|---|---|---|---|---|---|
| $^2$Co(II) | *4.1* | 27.1 | 27.7 | 13.2 | 16.3 | 9.2 | 9.1 |
| $^4$Co(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^6$Co(II) | *61.2* | 38.8 | 36.3 | 47.5 | 49.5 | 57.1 | 54.9 |
| $^1$Co(III) | *20.0*[b] | 23.8[c] | 24.6[c] | 19.1[c] | 16.9[c] | 21.1[c] | 19.0[c] |
| $^3$Co(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^5$Co(III) | *17.0* | 0.5 | −1.5 | 5.7 | 4.3 | 15.4 | 13.8 |
| $^1$Cr(II) | *67.7*[b] | 76.4[d] | 68.1[d] | 92.4[d] | 82.6[d] | 80.7[d] | 70.2[d] |
| $^3$Cr(II) | *35.9*[e] | 15.2[e] | 11.4[e] | 33.2[e] | 28.5[e] | 35.8[e] | 30.7[e] |
| $^5$Cr(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^2$Cr(III) | *39.4*[e] | 42.7[e] | 41.3[e] | 26.5[e] | 42.4[e] | 41.8[e] | 39.2[e] |
| $^4$Cr(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^1$Fe(II) | *48.2*[b] | 75.5[c] | 57.7[c] | 55.5[c] | 53.3[c] | 56.6[c] | 47.9[c] |
| $^3$Fe(II) | *19.1* | 41.0 | 42.0 | 22.9 | 21.5 | 27.1 | 19.3 |
| $^5$Fe(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^2$Fe(III) | *38.1*[e] | 41.1[e] | 41.8[e] | 35.9[e] | 35.7[e] | 36.1[e] | 35.6[e] |
| $^4$Fe(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^6$Fe(III) | *8.8* | −3.0 | −5.3 | −0.1 | −0.7 | 7.4 | 6.3 |
| $^2$Mn(II) | *70.9*[e] | 82.2[d] | 82.2[d] | 80.3[d] | 77.6[d] | 83.3[d] | 72.3[d] |
| $^4$Mn(II) | *26.3* | 21.2[e] | 19.1[e] | 28.6[e] | 26.0[e] | 27.0[e] | 24.8[e] |
| $^6$Mn(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^1$Mn(III) | *77.7*[b] | 111.4[d] | 107.7[d] | 106.8[d] | 101.0[d] | 72.8[d] | 77.2[d] |
| $^3$Mn(III) | *39.3* | 42.3[e] | 41.8[e] | 48.0[e] | 44.3[e] | 47.1[e] | 43.1[e] |
| $^5$Mn(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^1$Ni(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^3$Ni(II) | *4.5* | −17.4 | −17.6 | −7.0 | −6.3 | −0.1 | −0.2 |
| $^2$Ni(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^4$Ni(III) | *31.2* | 14.1 | 13.4 | 23.9 | 23.3 | 28.9 | 27.9 |
| $^1$Ti(II) | *9.5* | −2.0 | −4.0 | 7.9 | 3.9 | 10.5 | 6.2 |
| $^3$Ti(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^2$V(II) | *22.9*[e] | 10.0[e] | 6.8[e] | 23.6[e] | 18.1[e] | 25.1[e] | 19.9[e] |
| $^4$V(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^1$V(III) | *29.5* | 29.2 | 27.8 | 16.4 | 31.4 | 31.1 | 30.3 |
| $^3$V(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |

| M | *TZ2P*[a] | s3-21G | s3-21G* | s6-31G | s6-31G* | cs3-21G* | cs6-31G* |
|---|---|---|---|---|---|---|---|
| $^2$Co(II) | *4.1* | −3.1 | −1.8 | 1.0 | 2.6 | 10.6 | 5.2 |
| $^4$Co(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^6$Co(II) | *61.2* | 56.2 | 53.7 | 58.5 | 57.5 | 48.6 | 53.7 |
| $^1$Co(III) | *20.0*[b] | 16.9[c] | 17.2[c] | 20.6[c] | 18.8[c] | 17.6[c] | 18.9[c] |
| $^3$Co(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^5$Co(III) | *17.0* | 23.3 | 21.4 | 18.5 | 16.8 | 14.2 | 15.5 |
| $^1$Cr(II) | *67.7*[b] | 63.6[d] | 57.4[d] | 79.5[d] | 70.4[d] | 64.3[d] | 75.3[d] |
| $^3$Cr(II) | *35.9*[e] | 25.8[e] | 23.1[e] | 45.9[e] | 31.6[e] | 19.0[e] | 31.0[e] |
| $^5$Cr(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^2$Cr(III) | *39.4*[e] | 37.7[e] | 36.9[e] | 40.5[e] | 39.6[e] | 38.4[e] | 41.0[e] |
| $^4$Cr(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^1$Fe(II) | *48.2*[b] | 33.2[c] | 33.6[c] | 43.7[c] | 42.3[c] | 52.8[c] | 46.0[c] |
| $^3$Fe(II) | *19.1* | 10.0 | 10.4 | 16.6 | 15.8 | 25.9 | 18.3 |
| $^5$Fe(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^2$Fe(III) | *38.1*[e] | 32.4[e] | 32.6[e] | 36.0[e] | 35.6[e] | 31.8[e] | 34.8[e] |
| $^4$Fe(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^6$Fe(III) | *8.8* | 17.6 | 14.8 | 11.4 | 10.4 | 5.5 | 8.5 |
| $^2$Mn(II) | *70.9*[e] | 55.1[d] | 54.0[d] | 75.4[d] | 68.6[d] | 68.4[d] | 69.7[d] |
| $^4$Mn(II) | *26.3* | 13.4[e] | 13.0[e] | 23.0[e] | 21.1[e] | 18.9[e] | 23.9[e] |
| $^6$Mn(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |

**TABLE 4: Continued**

| M | *TZ2P*[a] | s3-21G | s3-21G* | s6-31G | s6-31G* | cs3-21G* | cs6-31G* |
|---|---|---|---|---|---|---|---|
| $^1$Mn(III) | *77.7*[b] | 65.5[d] | 79.5[d] | 73.3[d] | 67.6[d] | 83.5[d] | 66.0[d] |
| $^3$Mn(III) | *39.3* | 38.4[e] | 38.1[e] | 41.9[e] | 38.5[e] | 36.6[e] | 42.2[e] |
| $^5$Mn(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^1$Ni(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^3$Ni(II) | *4.5* | 8.1 | 7.4 | 4.9 | 4.7 | −2.5 | 1.1 |
| $^2$Ni(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^4$Ni(III) | *31.2* | 35.4 | 34.6 | 32.3 | 31.0 | 27.3 | 29.3 |
| $^1$Ti(II) | *9.5* | 7.8 | 6.0 | 12.4 | 8.4 | 1.6 | 7.0 |
| $^3$Ti(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^2$V(II) | *22.9*[e] | 16.8[e] | 14.3[e] | 24.4[e] | 19.4[e] | 11.8[e] | 19.6[e] |
| $^4$V(II) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |
| $^1$V(III) | *29.5* | 28.4 | 28.2 | 30.0 | 29.3 | 29.4 | 30.2 |
| $^3$V(III) | *0* | 0 | 0 | 0 | 0 | 0 | 0 |

[a] Reference data obtained with Slater-type orbital (TZ2P) basis set. [b] Obtained using augmented Roothaan−Hall (ARH) equations. [c] Obtained with spin-restricted Kohn−Sham. [d] Obtained using double spin contamination correction, see eq 2. [e] Obtained using spin contamination correction see eq 1.

We also investigated what happens when we mix the diffuse d-function with the small standard d-function instead of adding it as a third d-function (see Supporting Information). The results are presented in Tables 3 and 4 as cs3-21G* and cs6-31G* (for "correctly" spin-state-corrected basis sets). Although the results are similar to the corresponding s3-21G* and s6-31G* basis sets, they are not quite as good. This is especially evident for the cs6-31G* basis that has a maximum deviation that is ca. 1.5 times larger than that of s6-31G* and a MAD value of ca. 1.3 times larger. These larger deviations happen especially for the high spin states, which are typically placed at too low energy (see Table 4). Recently, we showed that this faulty behavior is typical for (small) GTO basis sets, and therefore we recommend the more appropriate s6-31G* basis set instead.

From the results shown above, it is now clear that one of the key ingredients for a correct description of the spin states of transition-metal complexes is the contraction scheme for the metal 3d orbitals. A [2d] contraction (as used in 6-31G*, m6-31G*, and cs6-31G*) is insufficient for correctly describing the relative spin-state energies, and at least a [3d] contraction (s6-31G*) is needed. This corroborates the results by Hay[16] who showed the same behavior for excitation energies.

**Accuracy of Geometries and Vibrational Frequencies.** We also investigated the effect of the basis set on the optimized geometries and vibrational frequencies (see Table 5). The results for original GTO basis sets show that a significant improvement is obtained when going from the 3-21G to the 6-31G basis set. These two basis sets have an MAD value when compared to the STO-TZ2P data of 0.048 Å, 95 cm$^{-1}$ (3-21G) and 0.017 Å, 59 cm$^{-1}$ (6-31G). Increasing the basis set size with the polarization functions actually makes the performance worse, that is, the MAD values for the geometry increase to 0.054 Å (3-21G*) and 0.021 Å (6-31G*), while the MAD for the frequencies of 3-21G* increases to 105 cm$^{-1}$. Only for the frequencies with 6-31G* do the polarization functions lead to a lowering of the MAD value, that is, compared to 6-31G the MAD drops from 59 to 37 cm$^{-1}$ (see Table 5). The modified 6-31G basis sets (m6-31G, m6-31G*) do not lead to an improvement for the geometries but instead lead to a larger MAD value for the geometries of 0.018 Å (m6-31G) and 0.019 Å (m6-31G*). For the frequencies, an improvement over the original 6-31G basis is observed for m6-31G (MAD value of 49 vs 59 cm$^{-1}$ for 6-31G), while the m6-31G* works as good as 6-31G* (MAD value of 37 cm$^{-1}$).

**TABLE 5: Mean Absolute Deviation (MAD) and Mean Deviation (MD) for the Accuracy of Geometries (Å) and Vibrational Frequencies (cm$^{-1}$) of Different GTO Basis Sets with Respect to Reference STO-TZ2P Data**

| | geometry | | vibrational frequencies | |
|---|---|---|---|---|
| basis set | MD | MAD | MD | MAD |
| 3-21G | −0.046 | 0.048 | 1 | 95 |
| 3-21G*[a] | −0.053 | 0.054 | −11 | 105 |
| 6-31G | 0.002 | 0.017 | 2 | 59 |
| 6-31G* | −0.012 | 0.021 | −3 | 37 |
| m6-31G | 0.013 | 0.018 | 15 | 49 |
| m6-31G* | −0.003 | 0.019 | 7 | 37 |
| | geometry | | vibrational frequencies | |
| basis set proposed here | MD | MAD | MD | MAD |
| s3-21G | −0.019 | 0.028 | 30 | 87 |
| s3-21G* | −0.028 | 0.031 | 20 | 89 |
| s6-31G | 0.011 | 0.015 | 15 | 50 |
| s6-31G* | −0.004 | 0.014 | 8 | 37 |
| cs3-21G*[b] | −0.030 | 0.033 | 12 | 83 |
| cs6-31G*[b] | −0.005 | 0.016 | 4 | 33 |

[a] The 3-21G* basis set contains a polarization f-function with exponent 0.8 similar to 6-31G*. [b] Diffuse d-function (see Table 1) mixed in with smallest standard d-function with weight 0.2 and 0.8 for diffuse and smallest function, respectively.

The spin-state-corrected basis sets show clear improvements over the original GTO basis sets. For the s3-21G and s3-21G* basis sets, an MAD value for the geometry is observed of 0.028 and 0.031 Å, respectively (see Table 5), that is, a reduction of ca. 40% compared to 3-21G and 3-21G*. The MAD value for the frequencies improves also somewhat to 87 and 89 cm$^{-1}$. Significant improvements are seen for the s6-31G and s6-31G* basis sets that show MAD values for the geometry of 0.015 and 0.014 Å, respectively. This is the only time that adding polarization functions actually helps to improve the MAD value for the geometries. Moreover, the value for the s6-31G* (0.014 Å) is the lowest MAD value for the small GTO basis sets studied here (see Table 5). The effect of the additional d-function in s6-31G and s6-31G* has almost no effect on the vibrational frequencies with MAD values that are equal (s6-31G*) to or slightly better (s6-31G) than the original GTO basis sets.

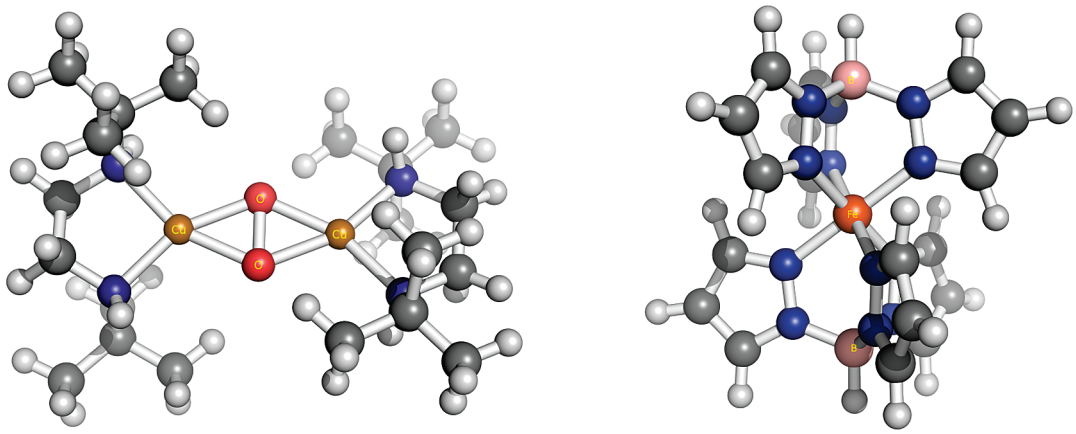


**Figure 2.** Transition-metal complexes used for timings of GTO basis sets, a 76-atom dicopper complex[22] ($\mu$-$\eta^2$:$\eta^2$-peroxodicopper(II) $Cu_2O_2(DBED)_2^{2+}$ [DBED = *N*,*N*′-di-*tert*-butylethylenediamine], left), and a 53-atom iron complex[20] ($Fe(T_b)_2$ [$T_b$ = trispyrazolylborate], right).

**TABLE 6: Time Needed (s) for 10 SCF Cycles with the Different GTO Basis Sets for the Dicopper and Iron Complexes (See Figure 2)**

| | $Cu_2O_2^{2+}$ complex | | Fe complex | |
|---|---|---|---|---|
| | time (s) | factor | time (s) | factor |
| 3-21G | 339.6 | | 59.0 | |
| 3-21G*[a] | 383.7 | *1.13*[b] | 64.4 | *1.09*[b] |
| 6-31G | 449.9 | *1.32*[b] | 73.5 | *1.25*[b] |
| 6-31G* | 756.1 | *1.68*[c] | 144.1 | *1.96*[c] |
| m6-31G | 464.7 | *1.03*[c] | 76.5 | *1.04*[c] |
| m6-31G* | 786.4 | *1.75*,[c] *1.04*[d] | 145.4 | *1.98*,[c] *1.01*[d] |
| basis set proposed here | $Cu_2O_2^{2+}$ complex | | Fe complex | |
| | time (s) | factor | time (s) | factor |
| s3-21G | 379.4 | *1.12*[b] | 62.4 | *1.06*[b] |
| s3-21G* | 428.2 | *1.26*[b],*1.12*[e] | 68.2 | *1.16*[b],*1.06*[e] |
| s6-31G | 503.7 | *1.12*[c] | 79.2 | *1.08*[c] |
| s6-31G* | 804.5 | *1.79*[c],*1.06*[d] | 150.6 | *2.05*[c],*1.05*[d] |
| cs3-21G*[f] | 405.4 | *1.19*[b],*1.06*[e] | 65.6 | *1.11*[b],*1.02*[e] |
| cs6-31G*[f] | 786.2 | *1.75*[c],*1.04*[d] | 146.8 | *2.00*[c],*1.02*[d] |

[a] The 3-21G* basis set contains a polarization f-function with exponent 0.8 similar to 6-31G*. [b] Factor with respect to 3-21G. [c] Factor with respect to 6-31G. [d] Factor with respect to 6-31G*. [e] Factor with respect to 3-21G*. [f] Diffuse d-function (see Table 1) mixed in with smallest standard d-function with weight 0.2 and 0.8 for diffuse and smallest function, respectively.

**Timing Results for the GTO Basis Sets.** To gain deeper insight into the computational costs for the spin-state corrections to the GTO basis sets, we have taken two prototypical transition-metal complexes from previous studies.[20,22] The two complexes are a 76-atom dicopper complex[22] ($\mu$-$\eta^2$:$\eta^2$-peroxodicopper(II) $Cu_2O_2(DBED)_2^{2+}$ [DBED = *N*,*N*′-di-*tert*-butylethylenediamine], Figure 2 left) and a 53-atom iron complex[20] ($Fe^{II}(T_b)_2$ [$T_b$ = trispyrazolylborate], Figure 2 right). Both were studied using a fine grid. To avoid ambiguity in the timings because of possibly a different number of SCF cycles needed to reach convergence, we simply took as measure for the computational cost the time needed to perform 10 SCF cycles. The timings for these complexes with the different basis sets are given in Table 6.

The largest differences in these timings result when introducing the polarization functions, for example, when going from 6-31G to 6-31G*, which leads to an increase of ca. 70−100% for the time needed for the 10 SCF cycles (see Table 6). This factor for the increase is virtually the same for the original Pople-type basis set (6-31G), the modified one (m6-31G), or the spin-state-corrected one (s6-31G). That is, the additional increase compared to, for example, 6-31G* for using the spin-state-corrected (s6-31G*) basis sets is very modest (6% or less, see Table 6). Given the enhanced performance and reliability of the spin-state-corrected GTO basis sets, this limited additional computational cost is a price worth paying.

## Conclusions

We report here simple and straightforward corrections to small Pople-type basis sets for first-row transition-metal atoms that correct their faulty behavior for spin-state splittings. The corrections were based on an even-tempered expansion of the 3d orbital space for the transition metal to make it more diffuse. This is an important correction especially for the newly constructed s6-31G* basis set. The latter basis set has small absolute deviations (1.1 kcal·mol$^{-1}$ for the spin-state splitting, 0.014 Å for geometries, and 37 cm$^{-1}$ for vibrational frequencies) from reliable reference data obtained with a Slater-type orbital basis set and correctly predicts the spin ground state for a set of 13 transition-metal complexes $MFHOH^{-1/0}$ with redox states II and III on the metal. This is a clear improvement on original Pople-type basis sets (6-31G, 6-31G*) and even on modified ones (m6-31G, m6-31G*), which fail in predicting the spin ground state of at least one transition-metal complex and which show larger deviations (1.6−4.1 kcal·mol$^{-1}$ for the spin-state splitting, 0.017−0.021 Å for the geometry, and 37−59 cm$^{-1}$ for the frequencies).

The spin-state corrections to the small 3-21G and 3-21G* basis sets (s3-21G, s3-21G*) drastically improve their results with a reduction of the deviations by half for the energy and geometry. That is, the MAD for the energy decreases from 7.1−7.8 kcal·mol$^{-1}$ (3-21G, 3-21G*) to 3.5−3.6 kcal·mol$^{-1}$ (s3-21G, s3-21G*), and the MAD for the geometry drops from 0.048−0.054 Å (3-21G, 3-21G*) to 0.028−0.031 Å (s3-21G, s3-21G*). Therefore, our previous notion that these small GTO basis sets lack a proper number of basis functions in the inner region is wrong. A simple spin-state correction largely corrects their faulty behavior as shown here.

An important factor for the performance of these small GTO basis sets was found to be the contraction scheme of the metal 3d orbitals, which has to be a [3d] contraction as already advocated in 1977 by Hay.[16] In contrast, m6-31G and m6-31G* keep the [2d] contraction of the 6-31G and 6-31G* basis sets, which explains their poorer results compared to the s6-31G* basis set. On the basis of these results, we recommend the use of the newly constructed s6-31G* for obtaining reliable spin-state splittings in GTO-based programs.

**Acknowledgment.** We thank the following organizations for financial support: the Ministerio de Ciencia e Innovación (MICINN, project numbers CTQ2008-03077/BQU and CTQ2008-06532/BQU), the DIUE of the Generalitat de Catalunya (project numbers 2009SGR637 and 2009SGR528), and a research grant for MG (MICINN). Support for the research of M. Solà was received through the prize "ICREA Academia" 2009 for excellence in research funded by the DIUE of the Generalitat de Catalunya.

**Supporting Information Available:** Example basis set description for iron with 6-31G*, m6-31G*, s6-31G*, and cs6-31G* basis sets. This material is available free of charge via the Internet at http://pubs.acs.org.

## References and Notes

(1) Swart, M.; Güell, M.; Solà, M. In *Quantum Biochemistry: Electronic structure and biological activity*; Matta, C. F., Ed.; Wiley: New York, 2010; Vol. 2, p 551.

(2) Meunier, B.; de Visser, S. P.; Shaik, S. *Chem. Rev.* **2004**, *104*, 3947.

(3) Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. *Chem. Rev.* **2005**, *105*, 2279.

(4) Swart, M.; Groenhof, A. R.; Ehlers, A. W.; Lammertsma, K. *J. Phys. Chem. A* **2004**, *108*, 5479.

(5) Swart, M. *J. Chem. Theory Comput.* **2008**, *4*, 2057.

(6) Swart, M.; Solà, M.; Bickelhaupt, F. M. *J. Chem. Phys.* **2009**, *131*, 094103.

(7) Güell, M.; Luis, J. M.; Solà, M.; Swart, M. *J. Phys. Chem. A* **2008**, *112*, 6384.

(8) Swart, M.; Ehlers, A. W.; Lammertsma, K. *Mol. Phys.* **2004**, *102*, 2467.

(9) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403.

(10) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(11) Bylaska, E. J.; de Jong, W. A.; Kowalski, K.; Straatsma, T. P.; Valiev, M.; Wang, D.; Apra, E.; Windus, T. L.; Hirata, S.; Hackler, M. T.; Zhao, Y.; Fan, P.-D.; Harrison, R. J.; Dupuis, M.; Smith, D. M. A.; Nieplocha, J.; Tipparaju, V.; Krishnan, M.; Auer, A. A.; Nooijen, M.; Brown, E.; Cisneros, G.; Fann, G. I.; Fruchtl, H.; Garza, J.; Hirao, K.; Kendall, R.; Nichols, J. A.; Tsemekhman, K.; Wolinski, K.; Anchell, J.; Bernholdt, D.; Borowski, P.; Clark, T.; Clerc, D.; Dachsel, H.; Deegan, M.; Dyall, K.; Elwood, D.; Glendening, E.; Gutowski, M.; Hess, A.; Jaffe, J.; Johnson, B.; Ju, J.; Kobayashi, R.; Kutteh, R.; Lin, Z.; Littlefield, R.; Long, X.; Meng, B.; Nakajima, T.; Niu, S.; Pollack, L.; Rosing, M.; Sandrone, G.; Stave, M.; Taylor, H.; Thomas, G.; van Lenthe, J.; Wong, A.; Zhang, Z. *NWChem 5.0*; Pacific Northwest National Laboratory: Richland, WA, 2006.

(12) Baerends, E. J.; Autschbach, J.; Bérces, A.; Berger, J. A.; Bickelhaupt, F. M.; Bo, C.; de Boeij, P. L.; Boerrigter, P. M.; Cavallo, L.; Chong, D. P.; Deng, L.; Dickson, R. M.; Ellis, D. E.; van Faassen, M.; Fan, L.; Fischer, T. H.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Groeneveld, J. A.; Gritsenko, O. V.; Grüning, M.; Harris, F. E.; van den Hoek, P.; Jacob, C. R.; Jacobsen, H.; Jensen, L.; Kadantsev, E. S.; van Kessel, G.; Klooster, R.; Kootstra, F.; van Lenthe, E.; McCormack, D. A.; Michalak, A.; Neugebauer, J.; Nicu, V. P.; Osinga, V. P.; Patchkovskii, S.; Philipsen, P. H. T.; Post, D.; Pye, C. C.; Ravenek, W.; Romaniello, P.; Ros, P.; Schipper, P. R. T.; Schreckenbach, G.; Snijders, J. G.; Solà, M.; Swart, M.; Swerhone, D.; te Velde, G.; Vernooijs, P.; Versluis, L.; Visscher, L.; Visser, O.; Wang, F.; Wesolowski, T. A.; van Wezenbeek, E. M.; Wiesenekker, G.; Wolff, S. K.; Woo, T. K.; Yakovlev, A. L.; Ziegler, T. *ADF 2007.01, SCM*; Amsterdam, The Netherlands, 2007.

(13) Swart, M.; Bickelhaupt, F. M. *J. Comput. Chem.* **2008**, *29*, 724.

(14) Swart, M.; Bickelhaupt, F. M. *Int. J. Quantum Chem.* **2006**, *106*, 2536.

(15) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. *J. Chem. Phys.* **1998**, *109*, 1223.

(16) Hay, P. J. *J. Chem. Phys.* **1977**, *66*, 4377.

(17) Martin, J.; Baker, J.; Pulay, P. *J. Comput. Chem.* **2009**, *30*, 881.

(18) Mitin, A. V.; Baker, J.; Pulay, P. *J. Chem. Phys.* **2003**, *118*, 7775.

(19) Solà, M.; Gonzalez, C.; Tonachini, G.; Schlegel, H. B. *Theor. Chim. Acta* **1990**, *77*, 281.

(20) Güell, M.; Solà, M.; Swart, M. *Polyhedron* **2010**, *29*, 84.

(21) Høst, S.; Olsen, J.; Jansík, B.; Thøgersen, L.; Jørgensen, P.; Helgaker, T. *J. Chem. Phys.* **2008**, *129*, 124106.

(22) Güell, M.; Luis, J. M.; Solà, M.; Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2009**, *14*, 229.

JP102712Z