JCTC Journal of Chemical Theory and Computation

The Basis Set Convergence of Spin−Spin Coupling Constants Calculated by Density Functional Methods

Frank Jensen*

Department of Chemistry, University of Southern Denmark,
DK-5230 Odense, Denmark

Received May 11, 2006

Abstract: The previously proposed polarization-consistent basis sets, optimized for density functional calculations, are evaluated for calculating indirect nuclear spin−spin coupling constants. The basis set limiting values can be obtained by performing a series of calculations with increasingly larger basis sets, but the convergence can be significantly improved by adding functions with large exponents. An accurate calculation of the Fermi-contact contribution requires the addition of tight *s* functions, while the paramagnetic spin−orbit contribution is sensitive to the presence of tight *p* functions. The spin-dipolar contribution requires the addition of *p*, *d*, and *f* functions. The optimal exponents for the tight functions can be obtained by optimizing the absolute sum of all contributions to the spin−spin coupling constant. On the basis of a series of test cases, we propose a standard set of tight *s*, *p*, *d*, and *f* functions to be added to the polarization-consistent basis sets. The resulting pcJ-*n* basis sets should be suitable for calculating spin−spin coupling constants with density functional methods.

I. Introduction

The use of theoretical predicted molecular properties for correlating with experimental information is becoming an increasingly popular method for obtaining information regarding molecular structure. While some properties, like vibrational frequencies and intensities, have been used for many years, spin-dependent properties have only come into focus more recently.[1] The theory for calculating nuclear magnetic resonance (NMR) spin−spin coupling constants is well-known[2] and was implemented in a semiempirical framework already in 1970.[3] In recent years, the methodology for calculating such properties with a variety of sophisticated methods has become generally available and a number of studies have appeared where the accuracy of various methods has been tested. Of prime concern from a practical point of view is the use of density functional theory,[4] because these methods are applicable to quite large systems at a favorable computational cost.[5]

NMR spin−spin coupling constants pose a challenging problem, as the results depend strongly on the quality of the wave function and the basis set. Hartree−Fock wave functions are in general unsuitable for calculating spin−spin coupling constants because of the presence of triplet instabilities, while correlated methods perform much better. Density functional methods are significantly better than Hartree−Fock and often provide an acceptable accuracy with only modest computational requirements. Recently, Keal and Tozer have proposed a family of functionals designed specifically for calculating nuclear shielding constants,[6] and one could imagine similar functionals optimized for spin-dependent properties.

The second component for performing calculations is the use of a basis set for expanding the molecular orbitals. While many standard basis sets are available for predicting, for example, geometries and relative energies, the calculation of molecular properties often requires more specialized basis sets. The basis set convergence for properties depending on the region of the wave function far from the nucleus can be improved by the addition of basis functions with small exponents,[7] while properties depending on the core region require the addition of functions with large exponents.[8−10] The calculation of indirect spin−spin coupling constants belongs to the latter category, and a nonrelativistic perturbational approach involves the four operators shown in eqs 1−4.[2,11]

* Corresponding author fax: +45-6615-8780; e-mail: frj@dou.dk.

$$\mathbf{H}^{DSO} = \frac{g_A g_B \mu_N^2}{2c^4} \frac{(\mathbf{r}_{iA}^t \mathbf{r}_{iB} - \mathbf{r}_{iB} \mathbf{r}_{iA}^t)}{\mathbf{r}_{iA}^3 \mathbf{r}_{iB}^3} \qquad (1)$$

$$\mathbf{H}^{PSO} = \frac{g_A \mu_N}{c^2} \frac{\mathbf{r}_{iA} \times \mathbf{p}_i}{r_{iA}^3} \qquad (2)$$

$$\mathbf{H}^{SD} = \frac{g_e \mu_B g_A \mu_N}{c^2} \frac{(\mathbf{r}_{iA}^t \mathbf{r}_{iA} - 3\mathbf{r}_{iA}\mathbf{r}_{iA}^t)}{r_{iA}^5} \mathbf{s}_i \qquad (3)$$

$$\mathbf{H}^{FC} = \frac{8\pi g_e \mu_B g_A \mu_N}{3c^2} \delta(\mathbf{r}_{iA}) \mathbf{s}_i \qquad (4)$$

Here, $\mathbf{r}_{iA}$ denoted the position vector between electron i and nucleus A; $\mu_{B/N}$ is the Bohr and nuclear magneton; $g_{e/A}$ is the electron and nuclear g factor, and $\mathbf{s}_i$ is the electron spin operator. The diamagnetic spin−orbit operator $\mathbf{H}^{DSO}$ gives a contribution which can be evaluated as an expectation value over the wave function, while the other three must be evaluated by response methods, formally given as a sum over all excited states in perturbation theory.

$$\mathbf{J} = \langle \Psi_0 | \mathbf{H}^{DSO} | \Psi_0 \rangle + \sum_k^{PSO,SD,FC} \sum_{n \neq 0} \frac{\langle \Psi_0 | \mathbf{H}^k | \Psi_n \rangle \langle \Psi_n | \mathbf{H}^k | \Psi_0 \rangle}{E_0 - E_n} \qquad (5)$$

The spin−spin coupling is a 3×3 tensor, but only the average isotropic component corresponding to $1/3$ of the trace of $\mathbf{J}$ is observed in solution, and we will consequently focus on this. It should be noted that the magnitude of coupling constants varies from essentially zero to several thousand hertz depending on the specific system and pairs of nuclei, and only part of the large variation is due to differences in the g factor for different elements and isotopes (e.g., 5.58 for ^{1}H and 0.40 for ^{14}N). Each of the four contributions may be either positive or negative, and the total coupling constant may consequently be either positive or negative, or close to zero because of the cancellation of terms.

From the form of the operators it is clear that the diamagnetic and paramagnetic spin−orbit ($\mathbf{H}^{DSO}$ and $\mathbf{H}^{PSO}$) and the spin-dipolar operators ($\mathbf{H}^{SD}$) primarily sample the inner region of the wave function, while the Fermi-contact term ($\mathbf{H}^{FC}$) evaluates the quality of the wave function *at the* nuclear positions. The representation of the excited states in the response part, on the other hand, is expected to be sensitive to the quality of the wave function in the region far from the nucleus. It is therefore clear that calculation of spin−spin coupling constants requires a careful consideration of the basis set, as the results are expected to be sensitive to functions with both large and small exponents. Indeed, previous calculations for correlated wave functions have shown that basis sets of at least triple-ζ-type quality augmented with both diffuse and tight functions are required to obtain accuracy results.[8,9,12] A similar study using density functional methods has also stressed the importance of adding tight *s* functions in order to evaluate the Fermi-contact contribution.[5,13] These studies have used the correlation consistent basis sets[14] where additional tight *s* functions were added by extending the regular basis set inward by a geometrical progression and decontraction of the basis sets. During the course of this work, Manninen and Vaara proposed to use basis sets complete to within a given threshold in a given exponent range for calculating magnetic properties.[15] After this paper was submitted, Deng et al. proposed to use the aug-cc-pVXZ basis sets augmented by tight *s* and *d* functions.[16]

We have previously proposed a series of basis sets optimized specifically for density functional methods, called polarization-consistent, which have been shown to improve the basis set convergence for several properties relative to other standard basis sets.[17] The notation is pc-*n*, where *n* indicates the level of polarization beyond the atomic system; that is, pc-0 is unpolarized, pc-1 is of double-ζ quality with a single polarization function, pc-2 is of triple-ζ quality with *d*- and *f*-type polarization functions, and so forth. In the present case, we will investigate the basis set convergence for spin−spin coupling constants with the polarization-consistent basis sets and suggest that the convergence can be improved by the addition of tight *s*, *p*, *d*, and *f* functions.

II. Computational Details

All calculations have been performed with the Dalton program package[18] and the B3LYP functional,[19] as this has been one of the more successful for predicting NMR properties.[11,20] Coupling constants refer to the most common isotope in all cases. Molecular geometries have been optimized at the B3LYP/6-31G(d,p) level. We emphasize that only the convergence with respect to basis set is investigated in the present case, and no attempt is made for comparing with experimental results. The latter requires attention to, for example, the molecular geometry and the quality of the exchange-correlation functional. It is expected that a very similar basis set convergence is observed for other functionals, and the present results should thus be valuable for probing the accuracy of various exchange-correlation functionals for predicting NMR spin−spin coupling constants.

III. Basis Set Convergence for Polarization-Consistent Basis Sets

The spin−spin coupling is a potential observable between all nuclei with a spin in a molecule, although only one-, two- and three-bond couplings are usually sufficiently large to be observed. The number of systems investigated by theoretical methods is relatively limited, and only a few guidelines are available for determining the importance of the above four contributions. One- and two-bond coupling constants (^1J and ^2J) are often dominated by the FC term, especially if one of the nuclei is hydrogen. Three-bond coupling constants (^3J), on the other hand, are often found to have substantial PSO and SD contributions. The DSO contribution is small in absolute magnitude (a few hertz) but may represent a sizable fraction for small coupling constants. These generalizations depend strongly on the system; the coupling constant involving fluorine and chlorine, for example, has large PSO and SD contributions for one-, two-, and three-bond coupling.[10,21] Furthermore, most of the investigations so far have concentrated on systems containing only elements from the first row in the periodic table. In the

Table 1. Molecular Systems Used in the Present Study

HF, H_2O, NH_3, CH_4
CO, N_2, F_2, SiS, P_2, Cl_2, F_2O, Cl_2S
HCN, C_2H_2, C_2H_4, C_2F_2, H_2CO, H_2CS

present case, we have used the systems in Table 1 for probing the performance. The basis set convergence for each of the four terms is found to be very similar for all of the systems, but only a few of the present systems have coupling constants where all four terms make a significant contribution. One of the exceptions is the F_2 system, which will be used as a representative case. We note that current exchange-correlation functionals have problems calculating accurate values for especially the FC term for fluorine-containing systems,[10,11,22] often being in error by a factor of 2. This problem is less relevant for the present case, where only the basis set convergence is investigated.

An initial exploration showed that the spin−spin coupling constant displays an erratic behavior when contracted basis sets are used, and all basis set optimizations have consequently been done using completely uncontracted basis sets. Furthermore, for some systems, the results obtained with the aug-pc-*n* basis sets, which include diffuse functions, are significantly different from those obtained with the regular pc-*n* basis sets. The difference diminishes rapidly as the size of the basis set increases. The results shown in Figure 1 indicate that both types of basis sets converge to a limiting value of ∼15 500 Hz for F_2, but the convergence is slow. It is therefore of interest to investigate to what extent the basis set can be improved in order to speed up the convergence.

The optimization of basis sets has typically employed a variational criterion, that is, minimizing the energy as a function of the basis function exponents. For properties other than the energy, there is no analogous variational principle, although a basis set converged value can be defined by the requirement that the addition of further functions does not cause changes within a specified threshold. In our initial exploration of extending basis sets by adding tight functions
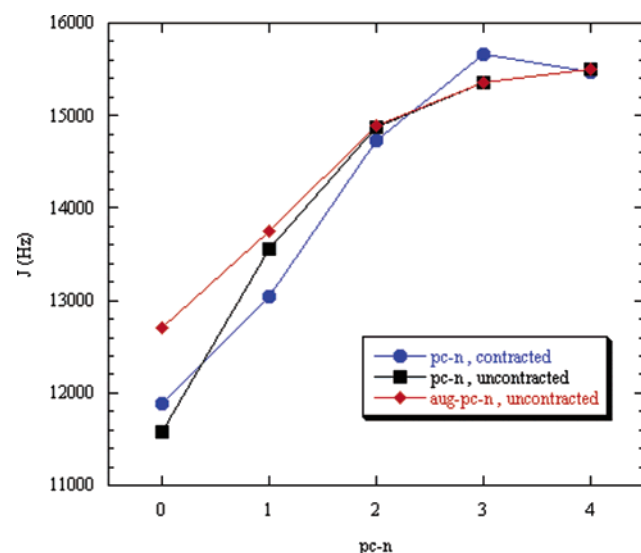
Figure 1. Coupling constant for F_2 calculated by the pc-*n* and aug-pc-*n* basis sets.

Table 2. DSO Contributions (Hz) to the Spin−Spin Coupling Constant in F_2[a]

basis	none	+s	+p	+d	+f
pc-0	3.41	3.41	3.04		
pc-1	2.53	2.53	2.39	2.36	
pc-2	2.20	2.20	2.17	2.14	2.13
pc-3	2.12	2.12	2.12	2.11	2.11
pc-4	2.11	2.11	2.11	2.11	2.11
aug-pc-0	3.30	3.30	2.99		
aug-pc-1	2.49	2.49	2.36	2.33	
aug-pc-2	2.20	2.20	2.17	2.14	2.13
aug-pc-3	2.12	2.12	2.12	2.11	2.11

[a] Notation +*x* indicates addition of a tight *x* function.

with exponents generated as an even-tempered sequence, we noted that the individual components of the coupling constant display a monotonic variation as more tight functions are added. Except for a few cases where the contribution is near zero and changes sign upon the addition of extra functions, this translates into a maximization of the absolute value of each of the contributions. We may consequently use the absolute value of each contribution as a parameter for optimizing the exponents of the added functions and to assess the importance of each type of function. For probing the sensitivity of the total coupling constant, which may contain both positive and negative contributions, we have used the quantity obtained by summation of the absolute values of each of the four contributions and used this in a variational sense to optimize the exponents of added functions.

The effect of the basis set on the coupling constant may conceptually be divided into three categories: (1) improving the representation of the wave function (density), (2) improving the representation of the operators, that is, improving the matrix elements in the nominator of the response part in eq 5, and (3) improving the excitation energies in the denominator of the response part in eq 5.

The pc-*n* basis sets have been designed to systematically improve category 1, which is the primary contributor to energetic quantities. On the basis of other work, it is expected that the addition of diffuse functions (aug-pc-*n* basis sets) will improve the performance for category 3. Our primary concern in the present case is therefore to add functions tailored to the representation of the operators in eqs 1−4. The basis set limiting value of the spin−spin coupling constant in F_2 is 15 556 Hz with the B3LYP functional, and we will in the following establish the requirements for converging the value to within ∼1 Hz of the limiting value.

A. The Diamagnetic Spin−Orbit Contribution. The DSO contribution is calculated as an expectation value over the unperturbed wave function and is in the employed implementation calculated by a grid approach.[10] It is therefore expected that this contribution is insensitive to the addition of tight functions, as their influence is only indirect by changes in the density. The calculated values for F_2 are given in Table 2, and the convergence is illustrated in Figure 2. The basis set limiting value is 2.1 Hz, and both the pc-*n* and aug-pc-*n* basis sets display a smooth convergence toward the limiting value. The convergence can be slightly improved by adding a tight *p* function, while the augmentation with diffuse functions has only a marginal effect (Table 2). The
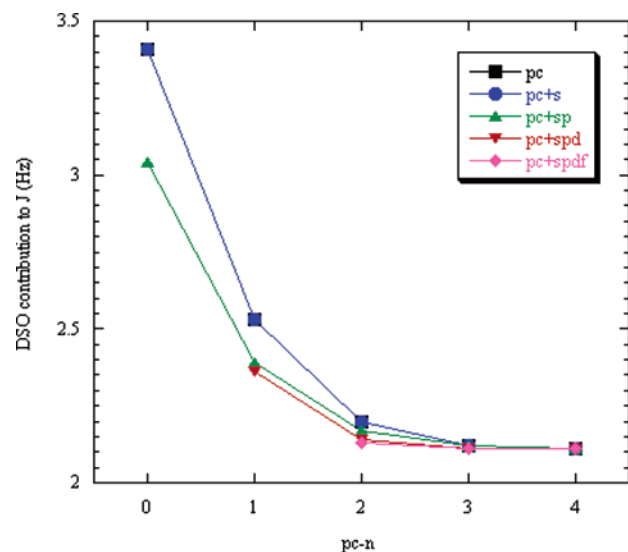
Convergence of Spin−Spin Coupling Constants

J. Chem. Theory Comput., Vol. 2, No. 5, 2006 **1363**

Figure 2. DSO contribution to the coupling constant for F_2 calculated by the pc-*n* basis sets augmented with tight functions.

Table 3. PSO Contributions (Hz) to the Spin−Spin Coupling Constant in $F_2{}^a$

basis	none	+s	+p	+d	+f
pc-0	7032	7033	9497		
pc-1	8542	8542	9569	9565	
pc-2	9245	9245	9418	9416	9409
pc-3	9399	9399	9417	9417	9416
pc-4	9409	9409	9414	9414	9414
aug-pc-0	7745	7746	10 483		
aug-pc-1	8623	8623	9661	9658	
aug-pc-2	9245	9245	9419	9416	9409
aug-pc-3	9395	9395	9414	9413	9413

a Notation +*x* indicates addition of a tight *x* function.

effect by addition of a tight *p* function is largest for the pc-0 and pc-1 basis sets, but because the DSO term is calculated as an expectation value, the change as a function of the basis set quality is due to improvements in representing the electronic density. The effect of the tight *p* function for the two smaller basis sets is therefore mainly due to inadequacies in the underlying pc-*n* basis set.

B. The Paramagnetic Spin−orbit Contribution. The PSO operator provides the largest contribution to the coupling constant in F_2, and the calculated values are given in Table 3 with the convergence illustrated in Figure 3. The addition of a tight *s* function has no effect, while the addition of a tight *p* function drastically improves the performance. For the pc-0 basis set, the change by adding a single tight *p* function is ∼2500 Hz, and the corresponding value for the pc-1 basis set is ∼1000 Hz. The addition of more tight *p* functions leads to further increases by ∼10%, that is, ∼250 and ∼100 Hz, respectively, but for the larger basis sets, the changes due to a second *p* function become negligible. Because the addition of just one tight *p* function slightly overestimates the basis set limiting value (Figure 3), this indicates that the addition of multiple tight *p* functions will be counterproductive, as it would cause the calculated value to move away from the limiting value, at the expense of
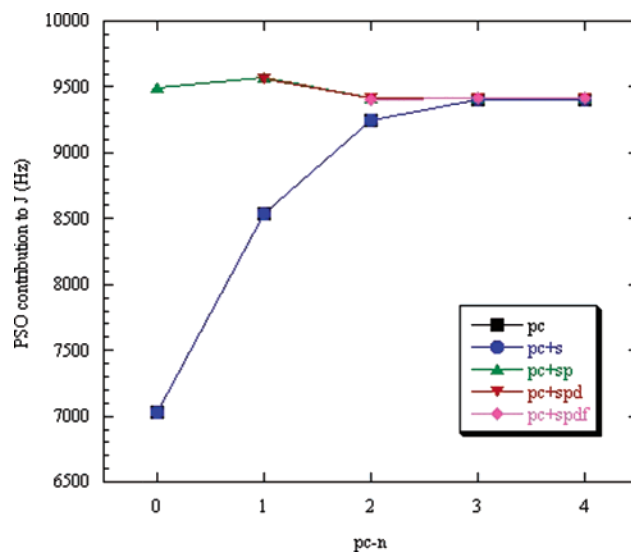

Figure 3. PSO contribution to the coupling constant for F_2 calculated by the pc-*n* basis sets augmented with tight functions.

Table 4. SD Contributions (Hz) to the Spin−Spin Coupling Constant in $F_2{}^a$

basis	none	+s	+p	+d	+f
pc-0	3773	3774	5172		
pc-1	4805	4805	5406	5691	
pc-2	5426	5426	5527	5800	5926
pc-3	5738	5738	5749	5873	5927
pc-4	5865	5865	5867	5873	5911
aug-pc-0	4341	4341	5969		
aug-pc-1	4870	4870	5481	5770	
aug-pc-2	5428	5428	5530	5803	5929
aug-pc-3	5737	5737	5748	5872	5926

a Notation +*x* indicates addition of a tight *x* function.

increasing the computational time. Higher angular momentum functions (*d* and *f* functions) have very marginal effects (i.e., a few hertz at most). Augmentation with diffuse functions has a large effect for the pc-0 basis set, but as seen in Table 3, the effect rapidly diminishes as the quality of the underlying basis improves.

C. The Spin Dipole Contribution. The SD operator provides the second largest contribution to the coupling constant in F_2, and the calculated values by adding tight functions are given in Table 4 and shown in Figure 4. In contrast to the PSO contribution, the SD contribution is sensitive to the addition of *p*, *d*, and *f* functions. Again, the majority of the effect (90+%) is obtained by adding a single tight function. The effect of adding a tight *g* function to the pc-3 basis set is only a fraction of a hertz. It should be noted that *removal* of the *g* function inherent in the pc-3 basis set causes changes of ∼10 Hz for the total coupling constant; that is, *g* functions are somewhat important, but the function already present in the pc-3 basis set is sufficient for converging the calculated value. The corresponding effect for the *h* function present in the pc-4 basis set is ∼1 Hz. Augmentation with diffuse functions (Table 4) again has a significant effect for the pc-0 and pc-1 basis sets but not for the larger basis sets.
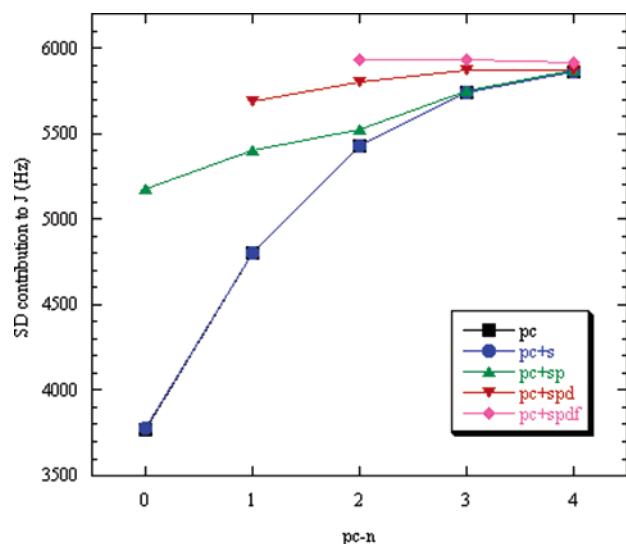
Figure 4. SD contribution to the coupling constant for F_2 calculated by the pc-*n* basis sets augmented with tight functions.

Table 5. FC Contributions (Hz) to the Spin−Spin Coupling Constant in F_2[a]

basis	none	+1s	+2s	+3s	+4s	+5s	+p	+d	+f
pc-0	774	986	1038	1053	1058	1059	1119		
pc-1	201	237	246	248	249	249	219	219	
pc-2	203	216	219	220	220		220	220	224
pc-3	226	230	231	231			231	231	232
pc-4	227	229	229				229	229	229
aug-pc-0	610	781	823	834	838	840	810		
aug-pc-1	256	302	313	316	317	317	288	288	
aug-pc-2	214	228	231	231			231	231	235
aug-pc-3	223	227	228	228			228	228	229

[a] Notation +*nx* indicates addition of *n* tight *x* function.

D. The Fermi-Contact Contribution. The FC operator contains a δ operator at the nuclear positions, and it is known that calculation of this term requires the addition of *s*-type functions with large exponents. In other work, it has been suggested that the addition of four to six functions with exponents generated by an even-tempered extension should be used for generating accurate results.[8,12,13,16] Higher-angular-momentum functions have nodes at the nuclear position and can only affect the FC term indirectly; that is, the density at a given nucleus can have contributions from functions centered at neighboring nuclei. Table 5 and Figure 5 show how the FC contribution converges upon the addition of *s* functions with optimized exponents. From Figure 5, it is clear that the first tight *s* function provides ∼70% of the total effect possible by saturating the function space with tight *s* functions. It is furthermore clear that the number of *s* functions required to reach the converged value decreases as the size of the underlying basis is increased. This is understandable as the larger basis sets already contain *s* functions with large exponents and, thus, have fewer require-ments for additional tight functions. Indeed, when examining the absolute value of the exponents, it is clear that the FC contribution is converged once the largest exponent reaches a value of ∼60×10^6 for F_2. For an optimum set of exponents, this requires five *s* functions for the pc-0 basis
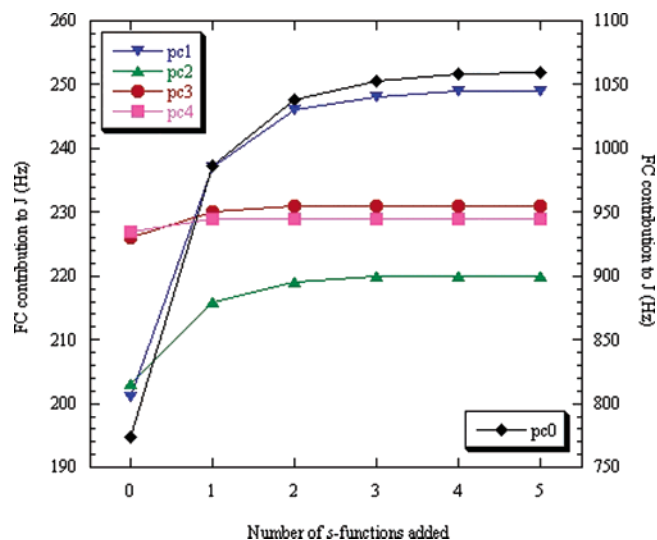

Figure 5. FC contribution to the coupling constant for F_2 as a function of added *s* functions with large exponents. The values on the left-hand axis refer to the pc-1, -2, -3, and -4 results; the values on the right-hand axis refer to the pc-0 results.

set, four for pc-1, three for pc-2, two for pc-3, and one for the pc-4 basis set.

While the addition of one tight *s* function clearly is desirable, the addition of multiple functions requires some considerations. Adding, for example, five tight *s* functions to the pc-0 basis set in order to converge the FC contribution will double the number of *s* functions and, thereby, make the number of *s* functions similar to that of the pc-2 basis set. The other contributions to the coupling constant, however, are relatively poorly represented with the pc-0 basis set. Adding many tight *s* functions in order to converge the FC contribution will thus not necessarily improve the total coupling constant, as the error in the other contributions may be dominating. Even for cases where the FC contribution dominates the total coupling constant, the representation of the density and excitation part of the response calculation is not significantly improved by the addition of tight *s* functions, and achieving the limiting FC contribution for a given quality of the underlying basis set is not necessarily a worthy goal. For the pc-1 basis set, for example, the changes by adding more than two *s* functions are only 3 Hz (Table 5), which should be compared to the error of 20 Hz relative to the basis set limit. Furthermore, the addition of several tight *s* functions increases the computational time, and from Figure 5, it is clear that this is a case of rapidly diminishing return. We suggest as a compromise between balance in the accuracy of all four terms, computational efficiency, and inherent accuracy of the underlying basis set, that two tight *s* functions should be added to the pc-0, pc-1, pc-2, and pc-3 basis sets but only one additional *s* function for the pc-4, as this essentially saturates the *s*-function space.

Table 5 and Figure 6 (pc-0 results are not show, as they clearly are of much lower accuracy) show that the FC contribution in addition to *s* functions is sensitive to the addition of a tight *p* function for the pc-0 and pc-1 basis sets. Analogous to the DSO contribution, this is most likely due to inadequacies in the underlying basis sets for repre-
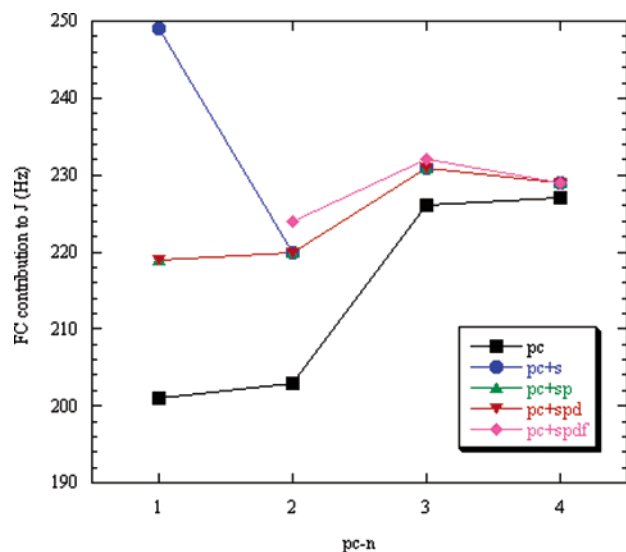
Figure 6. FC contribution to the coupling constant for F_2 calculated by the pc-*n* basis sets augmented with tight functions.

senting the density. Table 5 also shows that the FC contribution is sensitive to augmentation with diffuse functions, although the effect rapidly diminishes as the basis set quality is improved. It can also be noted that the presence of diffuse functions to some extent is counterproductive for this system, as the value moves further away from the limiting value.

IV. Designing Basis Sets for Calculating Spin−Spin Coupling Constants

The analysis in the previous section strongly indicates that a substantial improvement in the basis set convergence can be obtained by the addition of functions up to *f*-type with larger exponents than present in the standard basis sets. An explicit optimization of exponents for each specific system is impractical, and a procedure for selecting exponents for the additional tight functions is therefore required.

Previous work has recommended the addition of four tight *s* functions to basis sets of triple-ζ quality (e.g., pc-2) to converge the FC contribution.[8,12] The exponents of these functions were generated by multiplying the largest exponent with the ratio between the two innermost functions, that is, an even-tempered sequence. From our previous experience with basis set optimization, it is clear that this is unlikely to be optimum, as the ratio between exponents in a fully optimized basis set increases with the exponent values in the core region; that is, the spacing between exponents is small in the valence region and becomes larger as the functions extend into the core region. Furthermore, it is noticeable that the ratio between the two innermost *s* functions is a near constant value of 6.6 for all elements and basis sets ranging from pc-0 to pc-4. The same regularity is observed for the optimum exponents determined by optimizing the FC contribution for a variety of systems and basis sets, and suitable average values are given in Table 6. The exponent for a single tight *s* function should thus be generated by multiplying the innermost *s* function by a factor of 20; the exponents for two additional functions are obtained

Table 6. Recommended Exponent Ratios for Generating Tight Basis Functions

all elements					basis	hydrogen		1. row		2. row	
						p	d	d	f	d	f
1s	20.0				pc-1	9.0		20.0		180.0	
2s	12.5	25.0			pc-2	6.0	9.0	10.0	5.0	45.0	45.0
3s	9.0	12.5	25.0		pc-3	6.0	6.0	5.5	3.5	15	20.0
4s	8.0	9.0	12.5	25.0	pc-4	6.0	6.0	15.0	3.5	5.5	9.0
5s	8.0	8.0	9.0	12.5	25.0						
p	6.5										

by sequentially multiplying with factors of 12.5 and 25 and so forth. The exponent ratios in Table 6 allow a systematic and efficient convergence toward the basis set limit for the FC contribution, but as argued above, we recommend only the addition of two tight *s* functions, obtained by the sequential scaling of the innermost exponent by ratios of 12.5 and 25.0. We note that these two functions effectively cover the same exponent space as three functions generated by an even-tempered sequence.

Both the PSO and SD contributions are sensitive to the addition of tight *p*-type functions. An optimization of either the PSO or SD contribution separately indicated that the optimum exponent in both cases is close to that obtained by multiplying the innermost *p* function by a constant of 6.5. This validates the procedure used in section III where the exponents are optimized on the basis of the absolute sum of all four contributions. The ratio of 6.5 has been found to be remarkably constant for all of the basis sets and for a variety of systems. For the addition of two tight *p* functions, the corresponding ratios are 5.0 and 6.5, but the contribution from the second *p* function is only a few percent of the effect of the first function and is only significant for the smaller basis sets (pc-0 and pc-1). Arguing again for a balance in the accuracy of the four contributions, the quality of the underlying basis set, and the computational cost, the recommendation is to only extend the pc-*n* basis sets by one tight *p* function.

The SD contribution in addition requires augmentation with both tight *d*- and *f*-type functions. A determination of the optimum exponents for a selection of systems indicated that the optimum value is nearly constant for a given element, until the underlying pc-*n* basis set acquires exponents in the same range. This suggests that the SD contribution is sensitive to basis functions covering a rather specific distance from the nucleus. As expected the optimum value increases with the atomic number. The optimum exponents are again fairly close to those obtained by multiplying the innermost exponents by a fixed ratio, with the recommended values given in Table 6.

Our final recommendation is thus to add two tight *s* functions and one tight *p*, *d*, and *f* function to the pc-*n* basis sets, except that only one tight *s* function is added to the pc-4 basis set. The addition is of course conditioned on the presence of the corresponding type of function in the underlying basis set (e.g., *f* functions are not added to the pc-1 basis set). We will denote these pc-*n* basis sets augmented with tight functions as pcJ-*n*, where the J indicates that the additional functions have been optimized for
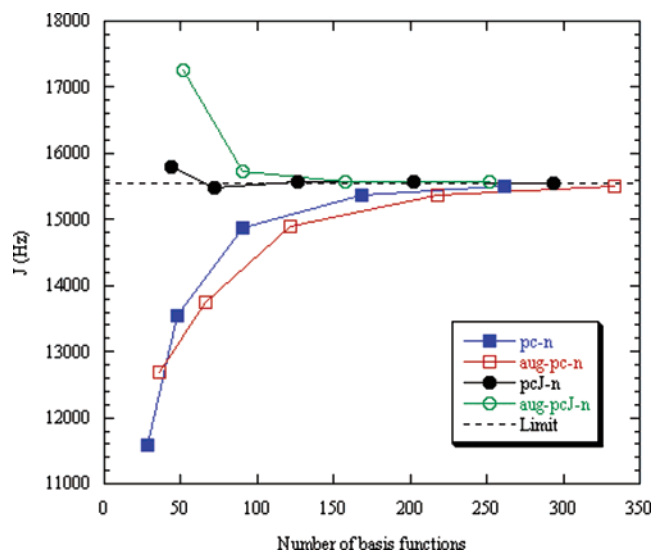
Figure 7. Coupling constant for F_2 calculated by the (uncontracted) pc-*n*, aug-pc-*n*, pcJ-*n*, and aug-pcJ-*n* basis sets.

calculation of the spin−spin coupling constants. Figure 7 shows the basis set convergence of the coupling constant in F_2 for the (uncontracted) pc-*n* and aug-pc-*n* basis sets and the two corresponding J-augmented ones, pcJ-*n* and aug-pcJ-*n*. The plot in Figure 7 has been done using the total number of basis functions as the variable, rather than the maximum angular momentum (Figure 1) to provide a more fair comparison. It is clear that the pcJ-*n* basis sets display a much improved performance, and even the double-ζ-quality basis pcJ-1 provides results close to the limiting value. It is also clear that the inclusion of diffuse functions is counterproductive for this specific system. The performance dependence on diffuse functions for a larger variety of systems will be quantified in section VI.

It is possible that the pcJ-*n* basis sets may also be useful for improving the basis set convergence for other properties depending on the atomic core region, but this clearly will require a careful calibration for each property.

V. Basis Set Contraction

The results in Figure 1 show that the standard contraction scheme is unsuitable for the calculation of spin−spin coupling constants. It should be noted that the F_2 system in this respect is relatively well-behaved, as significantly larger deviations between contracted and uncontracted basis sets have been observed for other systems. The sensitivity to contraction is not unexpected, as an increased flexibility in the core region is required for the calculation of spin−spin coupling constants. Although standard basis sets such as 6-311G** and cc-pVTZ have been tested for the calculation of spin−spin coupling constants,[16,21,23] the lack of tight functions and heavy contraction of the core region clearly makes such standard basis sets unsuitable.

The contraction of a basis set is always a balance between computational efficiency and a loss of accuracy, that is, what is the acceptable degradation of the results for a given gain in computational time. The acceptable loss of accuracy should be graded against the inherent error in the basis set relative to the basis set limiting result; that is, a relatively

large contraction error is acceptable if the underlying basis set has a large error relative to the limiting value anyway.

Previous work using the aug-cc-pVTZ basis set augmented with tight *s* functions recommended a partly uncontraction of the outer *s* function, while the contraction coefficients for the additional tight *s* functions were determined from molecular calculations.[8] It was shown that molecular contraction coefficients for one system could be used for others with little loss of accuracy, although the conclusion was based on only a few systems with similar bonding. The use of molecular information has in our opinion an inherent risk of biasing the results. The pc-*n* basis sets employ a general contraction using coefficients from atomic calculations, and we have shown that the contraction coefficient from one exchange-correlation functional can be used for other functionals with little loss of accuracy.[24] The maximum contraction is determined from the condition that the contraction error should be smaller than the inherent error of the uncontracted basis set relative to the basis set limit, and we have used the same approach in the present case.

Contraction of the inner *s* functions is expected to influence primarily the FC contribution. As the very tight functions describe the inner part of the 1s orbital, which is insensitive to the molecular environment, one might expect that this could be contracted with little loss of accuracy. On the other hand, the FC term is a δ function at a specific position, and even minor variations with the molecular species could potentially give substantial changes in the FC contribution. The results in Table 7 show that it is difficult to contract the *s* functions to any significant extent without destroying the accuracy of the uncontracted pcJ-*n* basis sets.

The PSO and SD contributions are insensitive to contraction of the *s*-function space but are sensitive to contraction of the *p*-function space. The contraction errors for the PSO and SD contributions are shown in Tables 8 and 9. It is again seen that only a small contraction of the inner functions is possible without compromising the accuracy. It is also seen that the PSO term is the most sensitive to contraction.

The results in Tables 7−9 illustrate that even the smaller of the pcJ-*n* basis sets (e.g., pcJ-1) are capable of providing results close to the basis set limiting values, and contraction errors must consequently be kept small. Given that the coupling constant depends primarily on the region near the nucleus, this means that it is difficult to contract the basis sets to a significant extent. It should be noted that the results in Tables 7−9 are specific for the F_2 system, but we have performed an analogous contraction analysis on the H_2, CO, C_2H_2, P_2, and SiS systems, with similar results. On the basis of these analyses, we performed calculations for all of the systems in Table 1 with a selection of contraction schemes, and the final recommended contraction schemes are given in Table 10.[25] The basis sets for the second row elements can be contracted somewhat harder than for the first row elements, presumably reflecting that the core region for second row elements is less affected by the molecular environment than for first row elements. The recommended contraction schemes are biased toward computational efficiency; that is, the contraction errors are comparable to or sometimes exceed the error of the uncontracted basis set

Table 7. Contraction Errors (Hz) for the FC Term in the F_2 Molecule[a]

basis	uncontracted	3s	4s	5s	6s	7s	8s	9s	10s	11s	12s	13s	14s	15s	16s	17s	18s	19s
pcJ-0	7s4p (866)	−243	92	46	1	0												
pcJ-1	9s5p (−13)		−515	71	−31	4	−1	0										
pcJ-2	12s7p (−6)			31	−37	20	−11	4	1	0								
pcJ-3	16s10p (3)					16	14	5	−5	3	−2	1	0	0	0			
pcJ-4	19s12p (0)								16	−13	9	−7	4	−4	1	−2	0	0

[a] Value in parentheses is the error of the uncontracted basis set relative to the pcJ-4 result.

Table 8. Contraction Errors (Hz) for the PSO Term in the F_2 Molecule[a]

basis	uncontracted	2p	3p	4p	5p	6p	7p	8p	9p	10p	11p	12p
pcJ-0	7s4p (76)	−201	6	0								
pcJ-1	9s5p (149)	−146	−45	−2	0							
pcJ-2	12s7p (−5)		−65	−20	−3	1	0					
pcJ-3	16s10p (2)			−70	−16	−8	−1	0	0	0		
pcJ-4	19s12p (0)					−9	−6	0	1	0	0	0

[a] Value in parentheses is the error of the uncontracted basis set relative to the pcJ-4 result.

Table 9. Contraction Errors (Hz) for the SD Term in the F_2 Molecule[a]

basis	uncontracted	2p	3p	4p	5p	6p	7p	8p	9p	10p	11p	12p
pcJ-0	7s4p (−743)	−134	2	0								
pcJ-1	9s5p (−227)	−149	−35	3	0							
pcJ-2	12s7p (16)		−83	−19	2	0	0					
pcJ-3	16s10p (16)			−76	−18	−1	1	0	0	0		
pcJ-4	19s12p (0)					−10	1	0	0	0	0	0

[a] Value in parentheses is the error of the uncontracted basis set relative to the pcJ-4 result.

Table 10. Recommended Contraction Schemes for the pcJ-n Basis Sets

basis	hydrogen		1. row		2. row	
	uncontracted	contracted	uncontracted	contracted	uncontracted	contracted
pcJ-0	5s	3s	7s4p	4s3p	10s7p	5s4p
pcJ-1	6s2p	4s	9s5p2d	5s4p	13s9p2d	6s5p
pcJ-2	8s3p2d	5s	12s7p3d2f	7s5p	15s11p3d2f	8s6p
pcJ-3	11s5p3d1f	8s	16s10p5d3f1g	10s8p	19s14p5d3f1g	11s10p
pcJ-4	12s7p4d2f1g	10s	19s12p7d4f2g1h	15s10p	22s17p7d4f2g1h	16s12p

relative to the basis set limit. It is, however, much easier to decontract a given basis set than to increase the contraction level, and the pcJ-n basis sets can easily be decontracted to reduce the contraction error if required by the specific system. It may be argued that the computational savings by contraction of the pcJ-3 and pcJ-4 basis sets are small, as the computational time tends to be dominated by the many polarization functions, and for benchmark calculations, one should consider using these basis sets in their uncontracted forms.

Although the requirement of using only a low contraction level will significantly increase the computational requirements, there are a few redeeming quantities to be considered. First, the pcJ-n basis sets already provide quite reasonable values for pcJ-1 and pcJ-2, which are applicable for a variety of systems. Furthermore, because the coupling constant to a large extent is determined by the atomic region, it is possible that the pcJ-n basis sets can be used as local dense basis sets for the nuclei between which the spin−spin coupling is desired, and regular (contracted) pc-n basis sets can be used for spectator atoms.[26] How well this will work in practice will require a calibration study and is outside the scope of the present work. For specific classes of coupling constants,

one may also consider pruning the pcJ-n basis sets selectively; for example, tight p, d, and f functions are not required for coupling constants that are dominated by the FC contribution, and tight d and f functions are only required if the SD contribution is a large component.

VI. Basis Set Convergence for a Larger Test Set

We have examined the performance of the pcJ-n and aug-pcJ-n families of basis sets for the systems in Table 1. Because coupling constants vary in magnitude by several thousand hertz, the error at a given level is reported as a percent deviation from the basis set limiting value for coupling constants larger than 10 Hz, as the percent deviation for small absolute values displays large nonsignificant fluctuations. The limiting value has in all cases been taken as the (uncontracted) aug-pcJ-4 result, and the convergence behavior indicates that this value is converged to at least 0.1%.

Table 11 shows the percent-wise mean and maximum absolute deviations (MAD and MaxAD) for the different basis sets over 33 coupling constants for the systems in Table 1. The effect of diffuse functions is system-dependent; that

Table 11. Percent-wise Mean and Maximum Absolute Deviations (MAD and MaxAD) Relative to the aug-pcJ-4 Results for Coupling Constants Larger than 10 Hz for the species in Table 1 (33 Data Points)

| | MAD | | | | | | MaxAD | | | | | |
| | uncontracted | | | | contracted | | uncontracted | | | | contracted | |
	pc-n	aug-pc-n	pcJ-n	aug-pcJ-n	pcJ-n	aug-pcJ-n	pc-n	aug-pc-n	pcJ-n	aug-pcJ-n	pcJ-n	aug-pcJ-n
pc-0	54.5	41.2	46.6	29.4	47.7	29.9	168.9	98.6	248.0	92.3	250.1	88.2
pc-1	18.5	18.6	4.5	3.2	6.5	4.8	34.1	32.4	19.9	11.7	20.8	25.8
pc-2	7.5	7.6	1.1	0.7	1.6	1.4	13.6	13.4	5.4	6.4	6.2	9.5
pc-3	2.2	2.1	0.3	0.1	0.4	0.8	4.2	3.8	3.1	0.7	3.4	2.8
pc-4	0.9	0.9	0.0	(0)	0.1	0.2	1.9	1.9	0.3	(0)	0.7	0.6

is, the HF, H_2O, F_2, CO, F_2O, C_2F_2, and P_2 systems are sensitive to augmentation with diffuse functions, while the effect for the other systems is very marginal. Furthermore, the effect of diffuse functions is largest for the smaller basis sets, that is, pcJ-0 and pcJ-1, while it rapidly diminishes as the underlying basis set is improved. The pcJ-0 basis set provides rather erratic results and is not recommended for general use. Given the relatively small improvement by adding diffuse functions to the pcJ-1 and pcJ-2 basis sets and the resulting increase in computational time, we do not generally recommend the inclusion of diffuse functions, which is in contrast to previous work using a wave function approach.[7]

The improved convergence indicated in Figure 7 for the F_2 system is also displayed by the MAD and MaxAD values in Table 11. It is encouraging that even the pcJ-1 basis set provides results within ~5% of the limiting value, and it will probably only in special cases be necessary to go beyond the pcJ-2 basis set in practical calculations. For very accurate work, the pcJ-3 and pcJ-4 basis sets should be able to provide the required accuracy. For the slightly smaller test set corresponding to removal of the systems containing Si, P, and Cl, the performance can be compared to the aug-cc-pVTZ-J basis set[7] which is comparable in size to pcJ-2. The MAD and MaxAD values for the aug-cc-pVTZ-J basis set are 1.8% and 8.0%, compared to the pcJ-2 values of 1.4% and 6.2%. Although this improvement is rather marginal, the present work has the advantage of defining a sequence of basis sets which allow a systematic improvement of the accuracy. It should also be noted that it is straightforward to define pcJ-n basis sets from the regular pc-n ones, when the latter are developed for more elements.

It should be noted that the basis set error is only one possible error component in a comparison with experimental values, as the reference geometry, vibrational averaging, solvent effects, and inadequacies in the exchange-correlation functional will need to be addressed in order to provide a direct comparison with experiments. The vibrational and environmental effects will typically change the coupling constant by ~5%,[5,27] while a typical error arising from the exchange-correlation functional is ~10%.[16] The present pcJ-n basis sets are optimized for density functional methods but may also be suitable for wave-function-based methods. It is likely, however, that more optimum basis sets for calculating spin−spin coupling constants with wave-function-based methods could be constructed from, for example, the

cc-pVXZ basis sets by augmenting them with tight functions determined along the principles used in the present work.

VII. Conclusions

The previously proposed polarization-consistent basis sets have been augmented with tight s, p, d, and f functions for improving the basis set convergence for calculating spin−spin coupling constants with density functional methods. The analysis shows that tight s functions are required for the Fermi-contact term, tight p functions are necessary for the paramagnetic spin−orbital contribution, while tight p, d, and f functions are required to converge the spin-dipolar contribution. The necessity of adding tight s functions has been noted by several other groups, but the requirement of tight p, d, and f functions for the noncontact terms is novel. The optimum exponents can be determined by a variational procedure based on the sum of the absolute value of all contributions. The optimum exponents display sufficient regularity from which general rules for assigning standard exponents for a given element can be derived. Using these rules, we propose a sequence of pcJ-n basis sets which is capable of converging the spin−spin coupling constant to within ~0.1% of the basis set limiting value. The basis set convergence of the pcJ-n basis sets is significantly improved relative to the regular basis sets, and the pcJ-1 and pcJ-2 basis sets should be efficient basis sets for calculating spin−spin coupling constants. While these basis sets by default contain tight s, p, d, and f functions, some of these functions may be omitted for specific cases. For coupling constants dominated by the Fermi-contact term, only the tight s functions are required, while a tight p function is only necessary if either the paramagnetic spin−orbital or spin-dipolar contribution is significant. The latter is often small, and in such cases, the tight d and f functions can be removed. Such basis set pruning, however, must be decided on a case-by-case basis.

Supporting Information Available: Exponents and contraction coefficients for the pcJ-n and aug-pcJ-n basis sets for the elements H, C, N, O, F, Si, P, S, and Cl. This material is available free of charge via the Internet at http://pubs.acs.org.

Convergence of Spin−Spin Coupling Constants

J. Chem. Theory Comput., Vol. 2, No. 5, 2006 **1369**

References

(1) Provasi, P. F.; Aucar, G. A.; Sanchez, M.; Alkorta, I.; Elguero, J.; Sauer, S. P. A. *J. Phys. Chem. A* **2005**, *109*, 6555. Gräfenstein, J.; Tuttle, T.; Cremer, D. *J. Phys. Chem. A* **2005**, *109*, 2325. Wu, A.; Cremer, D.; Auer, A. A.; Gauss, J. *J. Phys. Chem. A* **2002**, *106*, 657. Del Bene, J. E.; Perera, S. A.; Bartlett, R. J. *J. Am. Chem. Soc.* **2000**, *122*, 3560.

(2) Kirpekar, S.; Jensen, H. J. A.; Oddershede, J. *Chem. Phys.* **1994**, *188*, 171. Perera, S. A.; Nooijen, M.; Bartlett, R. J. *J. Chem. Phys.* **1996**, *104*, 3290. Helgaker, T.; Jaszunski, M.; Ruud, K. *Chem. Rev.* **1999**, *99*, 293. Melo, J. I.; Ruiz de Azua, M. C.; Peralta, J. E.; Scuseria, G. E. *J. Chem. Phys.* **2005**, *123*, 204112. Jensen, F. *Introduction to Computational Chemistry*; Wiley: New York, 1999. Kaupp, M.; Bühl, M.; Vladimir, M. G. *Calculation of NMR and ESR Parameters*; Wiley-VCH: New York, 2004. Sauer, S. P. A.; Packer, M. J. In *Computational Molecular Spectroscopy*; Jensen, P., Bunker, P. R., Eds.; Wiley: New York, 2000.

(3) Maciel, G. E.; McIver, J. W., Jr.; Ostlund, N. S.; Pople, J. A. *J. Am. Chem. Soc.* **1970**, *92*, 4151.

(4) Parr, R. G.; Yang, W. *Density Functional Theory*; Oxford University Press: New York, 1989. Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: New York, 2000.

(5) Lutnaes, O. B.; Ruden, T. A.; Helgaker, T. *Magn. Reson. Chem.* **2004**, *42*, S117.

(6) Keal, T. W.; Tozer, D. *J. Chem. Phys.* **2005**, *121*, 5654.

(7) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796. Jensen, F. *J. Chem. Phys.* **2002**, *117*, 9234.

(8) Provasi, P. F.; Aucar, G. A.; Sauer, S. P. A. *J. Chem. Phys.* **2001**, *115*, 1324.

(9) Enevoldsen, T.; Oddershede, J.; Sauer, S. P. A. *Theor. Chem. Acc.* **1988**, *100*, 275. Provasi, P. F.; Aucar, G. A.; Sauer, S. P. A. *Int. J. Mol. Sci.* **2003**, *4*, 231. Provasi, P. F.; Aucar, G. A.; Sauer, S. P. A. *J. Phys. Chem. A* **2004**, *108*, 5393.

(10) Barone, V.; Provasi, P. F.; Peralta, J. E.; Snyder, J. P.; Sauer, S. P. A.; Contreras, R. H. *J. Phys. Chem. A* **2003**, *107*, 4748.

(11) Helgaker, T.; Watson, M.; Handy, N. C. *J. Chem. Phys.* **2000**, *113*, 9402.

(12) Helgaker, T.; Jaszunski, M.; Ruud, K.; Gorska, A. *Theor. Chem. Acc.* **1998**, *99*, 175.

(13) Peralta, J. E.; Scuseria, G. E.; Cheeseman, J. R.; Frisch, M. J. *Chem. Phys. Lett.* **2003**, *375*, 452.

(14) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007. Wilson, A. K.; van Mourik, T.; Dunning, T. H., Jr. *J. Mol. Struct.* **1996**, *388*, 339.

(15) Manninen, P.; Vaara, J. *J. Comput. Chem.* **2006**, *27*, 434.

(16) Deng, W.; Chesseman, J. R.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 1028.

(17) Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113. Jensen, F. *J. Chem. Phys.* **2002**, *116*, 3502. Jensen, F. *J. Chem. Phys.* **2002**, *116*, 7372. Jensen, F. *J. Chem. Phys.* **2002**, *117*, 9234. Jensen, F. *J. Chem. Phys.* **2003**, *118*, 2459. Jensen, F.; Helgaker, T. *J. Chem. Phys.* **2004**, *121*, 3462.

(18) Helgaker, T.; Jensen, H. J. A.; Jørgensen, P.; Olsen, J.; Ruud, K.; Ågren, H.; Auer, A. A.; Bak, K. L.; Bakken, V.; Christiansen, O.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Enevoldsen, T.; Fernandez, B.; Hättig, C.; Hald, K.; Halkier, A.; Heiberg, H.; Hettema, H.; Jonsson, D.; Kirpekar, S.; Kobayashi, R.; Koch, H.; Mikkelsen, K. V.; Norman, P.; Packer, M. J.; Pedersen, T. B.; Ruden, T. A.; Sanchez, A.; Saue, T.; Sauer, S. P. A.; Schimmelpfenning, B.; Sylvester-Hvid, K. O.; Taylor, P. R.; Vahtras, O. *DALTON*, release 2.0; 2005.

(19) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(20) Maximoff, S. N.; Peralta, J. E.; Barone, V.; Scuseria, G. E. *J. Chem. Theory Comput.* **2005**, *1*, 541. Keal, T. W.; Helgaker, T.; Salek, P.; Tozer, D. J. *Chem. Phys. Lett.* **2006**, *425*, 163.

(21) Barone, V.; Peralta, J. E.; Contreras, R. H.; Snyder, J. P. *J. Phys. Chem. A* **2002**, *106*, 5607.

(22) Malkin, V. G.; Malkina, O. L.; Salahub, D. R. *Chem. Phys. Lett.* **1994**, *221*, 91. Malkin, O. L.; Salahub, D. R.; Malkin, V. G. *J. Chem. Phys.* **1996**, *105*, 8793. Peralta, J. E.; Barone, V.; Contreras, R. H.; Zaccari, D. G.; Snyder, J. P. *J. Am. Chem. Soc.* **2001**, *123*, 9162. Lantoo, P.; Vaara, J.; Helgaker, T. *J. Chem. Phys.* **2002**, *117*, 5998.

(23) Hansen, M. J.; Wendt, M. A.; Weinhold, F.; Farrar, T. C. *Mol. Phys.* **2002**, *100*, 2807.

(24) Jensen, F. *Chem. Phys. Lett.* **2005**, *402*, 510.

(25) All of the pcJ-*n* basis sets are available as Supporting Information.

(26) Sanchez, M.; Provasi, P. F.; Aucar, G. A.; Sauer, S. P. A. *Adv. Quantum Chem.* **2005**, *48*, 161.

(27) Mikkelsen, K. V.; Ruud, K.; Helgaker, T. *J. Comput. Chem.* **1999**, *20*, 1281. Ruden, T. A.; Helgaker, T.; Jaszunski, M. *Chem. Phys.* **2004**, *296*, 53.