

# XZP + 1d and XZP + 1d-DKH basis sets for second-row elements: application to CCSD(T) zero-point vibrational energy and atomization energy calculations

Cesar T. Campos · Francisco E. Jorge · Júlia M. A. Alves

Received: 20 January 2012 / Accepted: 19 March 2012 / Published online: 15 April 2012  
© Springer-Verlag 2012

**Abstract** Recently, segmented all-electron contracted double, triple, quadruple, quintuple, and sextuple zeta valence plus polarization function (XZP, X = D, T, Q, 5, and 6) basis sets for the elements from H to Ar were constructed for use in conjunction with nonrelativistic and Douglas–Kroll–Hess Hamiltonians. In this work, in order to obtain a better description of some molecular properties, the XZP sets for the second-row elements were augmented with high-exponent d “inner polarization functions,” which were optimized in the molecular environment at the second-order Møller–Plesset level. At the coupled cluster level of theory, the inclusion of tight d functions for these elements was found to be essential to improve the agreement between theoretical and experimental zero-point vibrational energies (ZPVEs) and atomization energies. For all of the molecules studied, the ZPVE errors were always smaller than 0.5 %. The atomization energies were also improved by applying corrections due to core/valence correlation and atomic spin-orbit effects. This led to estimates for the atomization energies of various compounds in the gaseous phase. The largest error (1.2 kcal mol<sup>-1</sup>) was found for SiH<sub>4</sub>.

**Keywords** XZP and XZP-DKH basis sets · Tight d functions · CCSD(T) method · Zero-point vibrational energy and atomization energy

## Introduction

The calculation of accurate bond energies or atomization energies has many important applications. While there are several approaches to this problem, the common thread is the determination of the equilibrium geometries and calculation of the zero-point vibrational energies (ZPVEs) at a modest level of theory, the calculation of the bond energies using an accurate treatment of the electron correlation problem [such as the coupled cluster with single, double, and perturbative triple excitation [1], CCSD(T)], and a correction to reach the basis set limit. The inherent accuracy of a correlated electronic structure method can only be assessed if the complete basis set (CBS) limit can be reliably determined. High-accuracy basis sets are thus desirable, as they provide estimates of the CBS limits.

Several basis sets that have been developed are capable of yielding very accurate results. Examples are the correlation-consistent polarized valence basis sets of double, triple, quadruple, quintuple, and sextuple (cc-pVXZ, X=D, T, Q, 5, and 6) zeta qualities developed by Dunning et al. [2–4] for H, B–Ne, and Al–Ar. These sets were augmented with diffuse functions in order to describe the long-range behavior of the wave function more accurately, and they were designated aug-cc-pVXZ (X=D, T, Q, and 5) [5]. Jensen [6, 7] used a hierarchy of polarization-consistent basis sets (pc-n, n=0, 1, 2, 3, and 4) to extrapolate to the Hartree–Fock (HF) and density functional theory (DFT) basis set limits. It should be noted that all of these sets were constructed from the general contraction scheme.

Jorge et al. presented segmented contracted double [8], triple and quadruple [9], quintuple [10], and sextuple [11] zeta valence quality plus polarization function (XZP, X=D,

C. T. Campos · F. E. Jorge (✉) · J. M. A. Alves  
Departamento de Física, Universidade Federal do Espírito Santo,  
29060-900 Vitória, ES, Brazil  
e-mail: jorge@cce.ufes.br

T, Q, 5, and 6, respectively) basis sets for the atoms from H to Ar. These have been successfully applied at the HF, DFT, second-order Møller–Plesset (MP2), and CCSD(T) levels of theory in calculations of various atomic and molecular properties [8–12].

There are a number of important effects that are not included in traditional molecular electronic structure calculations. Among the most important of these are relativistic effects, which include both spin–orbit and scalar relativistic effects. It was shown that the magnitude of the scalar relativistic corrections can be significant, even for first-row molecules, particularly for species which contain multiple halogen atoms (see [13] and references therein).

There are many ways to obtain scalar relativistic effects. Among these, we call attention to the Douglas–Kroll–Hess (DKH) approach [14–16], since it is known to account for most scalar relativistic effects [17].

The wave functions obtained using the DKH Hamiltonian—or any relativistic (corrected) Hamiltonian for that matter—will have a different radial behavior from that of the wave functions obtained from the Schrödinger Hamiltonian [17]. This difference can have a profound effect on the calculated result when nonrelativistic contracted basis sets are used. As contracted sets do not have enough flexibility in the core region to accommodate significant radial changes in the wave function, DKH calculations that introduce considerable radial changes (compared to the nonrelativistic radii) using nonrelativistic contracted basis sets are expected to yield poor results (see [17]). Therefore, in order to systematically incorporate the scalar relativistic effects from the DKH Hamiltonian, Jorge et al. [11, 18] recontracted the XZP (X=D, T, Q, 5, and 6) [8–11] basis sets for the elements from H to Ar. These sets were designated XZP-DKH.

The cc-pVXZ [2–4], pc-n [6, 7], and XZP [8–11] hierarchical sequences of basis sets have been used extensively in studies that include basis set extrapolation (see, e.g., [10–12, 19–25]). Extrapolation using the cc-pVXZ (X=T, Q, and 5) sets has been shown to yield accurate bond energies in most cases. However, Bauschlicher and Partridge [26] showed that missing tight functions on sulfur resulted in an inaccurate extrapolation for the atomization energy of SO<sub>2</sub>. That is, if a family of basis sets is deficient in some systematic way, extrapolation will not overcome the deficiency, and thus will not yield reliable results. This problem is not unique to SO<sub>2</sub>; it has also been found for several additional second-row compounds [19, 20, 27, 28].

In their study, Bauschlicher and Partridge [26] showed that adding a tight d function to the cc-pVXZ basis sets significantly improved the atomization energy of SO<sub>2</sub>. They suggested that adding two tight d functions to the cc-pVXZ sets would yield an accurate value. More recently, Martin [27] showed that the addition of diffuse functions was also

required. However, Martin did not use two tight d functions, but a series where a tight d was added to aug-cc-pVTZ, a 2d1f set of tight functions was added to aug-cc-pVQZ, and a 3d2f1g set was added to the aug-cc-pV5Z set. Using this series of basis sets in conjunction with a variable  $\alpha$  extrapolation approach, Martin found excellent agreement with experiment for the atomization energies of SO and SO<sub>2</sub>. The geometries and anharmonic force fields were also quite accurate. Martin optimized the geometries at the CCSD(T) level, which would be impractical for larger systems using these large basis sets. Martin did not include scalar relativistic effects and did not note that other extrapolation approaches did not agree well with the value obtained using the variable  $\alpha$  approach.

In order to improve geometry, atomization energy, and other properties of molecules containing second-row elements, we decided to add a tight d polarization function to the nonrelativistic and relativistic basis sets reported in [8–11, 18]; i.e., to generate XZP+1d and XZP+1d-DKH (X=D, T, Q, 5, and 6) basis sets for the elements from Na to Cl. To our knowledge, this represents the first time that XZP+1d-DKH basis sets have been generated for these elements. In the work described in the present paper, we focused mainly on atomization energies, using the highest-level ab initio method that can be routinely applied to small molecules with extended basis sets. At the CCSD(T) level of theory, the effect of including tight d functions for second-row elements when calculating ZPVE and atomization energy was determined. The atomization energy CBS limits were also estimated using an extrapolation scheme. For SO<sub>2</sub>, the effect of tight d functions on geometric parameters and harmonic vibrational frequencies was also examined.

#### Tight d functions for NaCl

The XZP+1d set was obtained from the corresponding XZP representation by adding a high-exponent d polarization function. For each atom, the value of the d exponent was optimized by using the minimum restricted open-shell MP2 energy criterion for various molecules, namely: Na<sub>2</sub>, NaH, NaF, NaCl, MgH, MgO, MgF, AlH, AlF<sub>3</sub>, SiH<sub>4</sub>, SiO, PH<sub>3</sub>, PN, H<sub>2</sub>S, SO<sub>2</sub>, ClF, and HCl. The optimized values of tight d polarization functions were then averaged to a single value for each second-row element. It should be mentioned here that the same exponent generated at the nonrelativistic level of theory was also added to the corresponding relativistic basis set. The final values of these exponents are displayed in Table 1. The calculations were carried out using the frozen core (1s, 2s, 2p for Na–Cl) approximation for electron correlation and a subroutine developed by our research group that optimizes the d exponent of a given atom in the molecular environment and makes use of the MP2 method implemented in the Gaussian 09 code [29]. Throughout the

**Table 1** Tight d functions added to the nonrelativistic and relativistic basis sets of Na-Cl

	DZP+1d; DZP+1d-DKH	TZP+1d; TZP+1d-DKH	QZP+1d; QZP+1d-DKH	5ZP+1d; 5ZP+1d-DKH	6ZP+1d; 6ZP+1d-DKH
Na	0.24785600	0.63587333	0.93076500	0.91297500	5.25000000
Mg	0.39573333	0.64578033	2.03600000	2.13290000	6.12500000
Al	0.55197499	1.10805000	1.79154699	2.19813500	6.00000000
Si	0.81907349	1.77670000	2.45030500	3.05374600	7.50000000
P	1.03001500	2.30430000	3.30574999	4.18487448	19.56250000
S	1.43538500	3.01725500	4.28275149	5.16724999	13.06250000
Cl	1.57549500	3.67014999	5.15899999	8.64484948	20.37500000

calculations, we employed spherical harmonic Gaussian-type functions.

### Computational details

All molecular calculations were performed with the Gaussian 09 program [29].

Four families of all-electron correlated basis sets were used in combination with the CCSD(T) wave function. The XZP and XZP+1d (X = D, T, Q, 5, and 6) basis sets, which were developed for use explicitly with a nonrelativistic Hamiltonian, contained the standard published contraction coefficients. On the other hand, the XZP-DKH and XZP+1d-DKH contracted basis sets were designed for use with the relativistic DKH2 Hamiltonian.

In the CCSD(T) calculations, we did not include the full correlation energy correction; we limited it to the valence atomic orbitals. This approximation is known as the frozen core for electron correlation, and it significantly reduces the computational time (mainly for molecules containing elements from the second row on) when highly accurate methods [e.g., CCSD(T)] and large basis sets (e.g., 6ZP) are used to evaluate various properties. Except in the geometric parameter and harmonic vibrational frequency calculations for SO<sub>2</sub>, experimental geometries [30] were used, namely: SiH<sub>4</sub> (1.4798 Å, T<sub>d</sub>), PH<sub>3</sub> (1.4200 Å, ∠HPH 93.345°), H<sub>2</sub>S (1.3356 Å, ∠HSH 92.12°), HCl (1.2746 Å), Na<sub>2</sub> (3.0789 Å), P<sub>2</sub> (1.8931 Å), AlF<sub>3</sub> (1.6544 Å, D<sub>3h</sub>), SiCl<sub>4</sub> (2.019 Å, T<sub>d</sub>), and SO<sub>2</sub> (1.4308 Å, ∠OSO 119.329°).

In this study, we used one of the most popular forms of extrapolation

$$Y_R(X) = Y_R(\infty) + AX^{-3} \quad (1)$$

which is the basis for the CBS model of Helgaker et al. [31, 32].

In Eq. 1,  $Y_R(X)$  is the property computed at the internuclear distance  $R$ , and  $X$  denotes the highest angular functions of the basis sets used in the extrapolation.  $Y_R(\infty)$  is the property value in the complete basis set limit.  $A$  is a fitting parameter that does not have any physical significance.

Since there are two unknown quantities in Eq. 1 [ $Y_R(\infty)$  and  $A$ ], at least two consecutive basis sets are needed for extrapolation. Restricting ourselves to this requirement, we have used two member groups (X-1, X) in our relativistic hierarchical sequence. From here on, the CBS limits obtained from the 5ZP- and 6ZP-DKH and 5ZP+1d- and 6ZP+1d-DKH results will be denoted CBS(5,6-DKH) and CBS(5,6+1d-DKH), respectively.

### Results and discussion

Ab initio harmonic vibrational frequencies are normally larger than the fundamentals observed experimentally [33]. The main reason for this disagreement is that anharmonicity effects are neglected in the theoretical treatment. Errors also arise because of the incomplete treatment of electron correlation and basis set truncation effects. To improve the agreement between the predicted and experimentally observed [30] frequencies, the computed harmonic frequencies are usually scaled for comparison. The determination of appropriate scaling factors to estimate experimental fundamental frequencies from theoretical harmonic frequencies has received considerable attention in the literature [34, 35]. Recently, scaling factors for fundamental vibrational frequencies and ZPVEs obtained from HF, MP2, and DFT/DZP and TZP harmonic frequencies were reported by Andrade et al. [36]. It should be mentioned here that some methods (see, e.g., [37–39] and references therein) of computing anharmonic vibrational frequencies have been developed and implemented in various software packages.

Table 2 displays theoretical and experimental [19] ZPVEs for the compounds studied here. As the differences among the MP2 and CCSD(T)/TZP ZPVEs do not exceed 0.43 %, the MP2 scaling factor (0.9775) estimated previously [36] was used to correct the CC results. The ground-vibrational-state atomization energies ( $D_0$ ) shown in Table 3 were then evaluated from the scaled CCSD(T)/TZP and CCSD(T)/TZP+1d ZPVEs. The former was used to correct the equilibrium dissociation energy obtained from the XZP and XZP-DKH (X=D, T, Q, 5, and 6) sets, whereas the

**Table 2** Theoretical and experimental zero-point vibrational energies (in kcal mol<sup>-1</sup>)

Molecule	MP2/TZP		MP2/TZP+1d		CCSD(T)/TZP		CCSD(T)/TZP+1d		Expt. <sup>b</sup>
	Unscaled	Scaled <sup>a</sup>	Unscaled	Scaled <sup>a</sup>	Unscaled	Scaled <sup>a</sup>	Unscaled	Scaled <sup>a</sup>	
SiH <sub>4</sub>	20.08	19.63	20.04	19.59	19.80	19.35	19.76	19.32	19.21
PH <sub>3</sub>	15.32	14.97	15.28	14.94	15.14	14.80	15.09	14.75	-
H <sub>2</sub> S	9.73	9.51	9.70	9.48	9.65	9.43	9.62	9.40	9.4
HCl	4.37	4.27	4.35	4.26	4.35	4.25	4.33	4.23	4.2
Na <sub>2</sub>	0.24	0.24	0.24	0.24	0.24	0.23	0.24	0.23	0.2
P <sub>2</sub>	1.14	1.11	1.12	1.10	1.16	1.14	1.15	1.12	1.1
AlF <sub>3</sub>	4.80	4.70	4.75	4.64	4.81	4.70	4.75	4.65	-
SiCl <sub>4</sub>	4.77	4.66	4.74	4.64	4.77	4.66	4.74	4.63	-
SO <sub>2</sub>	4.55	4.45	4.47	4.37	4.57	4.46	4.49	4.39	4.4

<sup>a</sup> Scale factor (0.9775) is from [36]. It was obtained from MP2/TZP zero-point vibrational energy nonrelativistic calculations

<sup>b</sup> From [19]

latter was used to correct the others. The effect of the tight d function on ZPVE is small (<1.7 %), but it always improves the agreement between theory and experiment (see Table 2). It should be noted that there is excellent agreement between the CCSD(T)/TZP+1d results and the experimental data [19] available in the literature.

Calculated and experimental [40, 41] atomization energies for nine molecules containing a second-row element are presented in Table 3.  $D_0$  differs from the equilibrium dissociation energy by the ZPVE.

A brief look at Table 3 offers some general trends. For both the nonrelativistic and relativistic hierarchical sequences of basis sets, one can observe that from the triple zeta basis sets results,  $D_0$  increases smoothly, and that the agreement with the experimental atomization energies improves as the size of the basis set increases. The convergence is faster for hydride molecules and Na<sub>2</sub>; in these cases, the 6ZP-DKH results can be considered to be good estimates of the CBS limits. On the other hand, SO<sub>2</sub> exhibited extremely slow convergence with respect to the original correlated basis set progression (XZP and XZP-DKH). As mentioned before, it has been the subject of high-level studies. Martin [27] has emphasized the importance of tight (i.e., large exponent) d and f polarization functions on sulfur, and he concluded that in order to obtain a reliable extrapolation to the one-particle basis set limit, the addition of this kind of function was essential. In fact, tight d functions were added for all second-row atoms in order to improve the basis set convergence in atomization energy calculations for various compounds (see [19]). Our XZP+1d and XZP+1d-DKH results (shown in Table 3) add credence to this, and justify the addition of the tight d function for second-row elements. The effect of this kind of function on  $D_0$  decreases with basis set size. For SO<sub>2</sub>, it goes from ~25 to 0.94 kcal mol<sup>-1</sup> when basis sets of double and sextuple zeta qualities are

used, respectively. Consequently, for all compounds studied, one can verify that both the XZP-DKH and the XZP+1d-DKH hierarchical sequences of basis sets converge to practically the same values; e.g., to 251.05 and 251.99 kcal mol<sup>-1</sup>, respectively, for SO<sub>2</sub>. The main difference between these sequences is that the XZP+1d-DKH one goes to the CBS limit faster; i.e., the difference between the atomization energies obtained with adjacent basis sets that belong to the latter hierarchical sequence is smaller. Except for hydrides, this can lead to very different extrapolation values. For example, for SO<sub>2</sub>, the difference between the CBS(5,6-DKH) and CBS(5,6+1d-DKH) results is 4.88 kcal mol<sup>-1</sup>. Even for Na<sub>2</sub>, the CBS(5,6+1d-DKH) atomization energy is about 6 % larger than the CBS(5,6-DKH) one.

To determine the magnitude of the scalar relativistic correction, the difference between relativistic and nonrelativistic atomization energies was computed. The scalar relativistic correction lowers the atomization energy relative to the energy obtained from nonrelativistic calculations. Additionally, the absolute value of the correction is typically less than 0.9 kcal mol<sup>-1</sup>, though there are two exceptions.

Scalar relativistic contributions are summarized in Table 3. Even though the effect of the high-exponent d polarization function on the scalar relativistic correction is small, we believe that the results obtained with the sets containing this function are more reliable, so we used them in the subsequent analysis.

We first consider basis set convergence for the CCSD(T) results. From Table 3, it is clear that from DZP+1d-DKH to TZP+1d-DKH and from TZP+1d-DKH to QZP-DKH there are abrupt variations in the value of the scalar relativistic effect, but a smooth convergence is observed from QZP+1d-DKH, and it can be assumed to be achieved at the 6ZP+1d-DKH level of theory (except for AlF<sub>3</sub> and SO<sub>2</sub>), since the

**Table 3** CCSD(T) atomization energies<sup>a</sup> of some polyatomic molecules. In the  $D_0$  calculations, the frozen core approximation for electron correlation was used, and the scalar relativistic effects [CCSD(T)/XZP-DKH – CCSD(T)/XZP or CCSD(T)/XZP + 1d-DKH – CCSD(T)/XZP + 1d] are given in parentheses

	SiH <sub>4</sub> ( <sup>1</sup> A <sub>1</sub> )	PH <sub>3</sub> ( <sup>1</sup> A <sub>1</sub> )	H <sub>2</sub> S ( <sup>1</sup> A <sub>1</sub> )	HCl ( <sup>1</sup> Σ <sup>+</sup> )	Na <sub>2</sub> ( <sup>1</sup> Σ <sub>g</sub> <sup>+</sup> )	P <sub>2</sub> ( <sup>1</sup> Σ <sub>g</sub> <sup>+</sup> )	AlF <sub>3</sub> ( <sup>1</sup> A <sub>1</sub> )	SiCl <sub>4</sub> ( <sup>1</sup> A <sub>1</sub> )	SO <sub>2</sub> ( <sup>1</sup> A <sub>1</sub> )
DZP <sup>b</sup>	282.44	204.02	157.27	93.65	15.65	88.71	363.61	325.95	181.57
TZP <sup>c</sup>	298.08	219.39	167.96	99.64	16.14	102.10	406.71	360.09	225.60
QZP <sup>e</sup>	302.72	223.75	171.62	101.71	16.29	108.68	418.05	373.07	239.56
5ZP <sup>d</sup>	303.73	225.05	172.66	102.43	16.30	111.41	421.25	378.52	245.63
6ZP <sup>e</sup>	304.82	226.24	173.51	102.87	-	113.00	424.12	382.61	251.88
DZP-DKH <sup>f</sup>	280.86 (-1.58)	202.87 (-1.15)	156.44 (-0.83)	93.25 (-0.40)	15.66 (0.01)	88.45 (-0.26)	362.16 (-1.45)	324.43 (-1.52)	180.90 (-0.67)
TZP-DKH <sup>f</sup>	297.36 (-0.27)	219.04 (-0.35)	167.50 (-0.46)	99.37 (-0.27)	16.09 (-0.05)	101.94 (-0.16)	405.46 (-1.25)	358.35 (-1.74)	224.78 (-0.82)
QZP-DKH <sup>f</sup>	302.05 (-0.67)	223.30 (-0.45)	171.23 (-0.39)	101.47 (-0.24)	16.30 (0.01)	108.42 (-0.26)	416.73 (-1.32)	371.42 (-1.65)	238.83 (-0.73)
5ZP-DKH <sup>g</sup>	303.06 (-0.67)	224.59 (-0.46)	172.27 (-0.39)	102.18 (-0.25)	16.31 (0.01)	111.15 (-0.26)	419.89 (-1.36)	376.86 (-1.66)	244.84 (-0.79)
6ZP-DKH <sup>g</sup>	304.14 (-0.68)	225.78 (-0.46)	173.12 (-0.39)	102.63 (-0.24)	-	112.73 (-0.27)	422.75 (-1.37)	380.92 (-1.69)	251.05 (-0.83)
CBS(5,6-DKH) <sup>h</sup>	305.62	227.41	174.29	103.25	16.32	114.90	426.68	386.50	259.58
CBS(5,6-DKH) <sup>i</sup>	(-0.69)	(-0.46)	(-0.39)	(-0.23)	(0.01)	(-0.29)	(-1.38)	(-1.73)	(-0.88)
ΔE <sub>CV</sub> <sup>j</sup>	-0.2	0.4	0.4	0.3	0.3	0.8	-0.38	0.31	1.0
ΔE <sub>SO</sub> <sup>k</sup>	-0.4	0.0	-0.6	-0.8	0.0	0.0	-1.36	-3.79	-1.0
6ZP-DKH+ΔE <sub>CV</sub> +ΔE <sub>SO</sub>	304.08	225.82	173.10	102.58	16.61	113.53	421.01	377.44	251.05
CBS(5,6-DKH)+ΔE <sub>CV</sub> +ΔE <sub>SO</sub>	305.56	227.45	174.27	103.20	16.62	115.70	424.94	383.02	259.58
DZP + 1d <sup>l</sup>	288.30	208.63	160.52	95.31	15.98	93.58	383.87	340.54	205.85
TZP + 1d <sup>l</sup>	299.86	220.95	169.22	100.32	16.28	103.91	410.83	365.10	235.96
QZP + 1d <sup>l</sup>	303.89	224.85	172.42	102.13	16.38	110.00	420.42	377.04	246.50
5ZP + 1d <sup>l</sup>	304.61	225.82	173.27	102.66	16.40	112.80	423.06	381.09	250.85
6ZP + 1d <sup>l</sup>	305.03	226.32	173.65	102.90	-	113.46	424.48	383.06	252.84
DZP + 1d - DKH <sup>l</sup>	286.81 (-1.49)	207.60 (-1.03)	159.79 (-0.73)	94.97 (-0.34)	15.99 (0.01)	93.30 (-0.28)	382.39 (-1.48)	338.93 (-1.61)	205.01 (-0.84)
TZP + 1d - DKH <sup>l</sup>	299.11 (-0.75)	220.57 (-0.38)	168.74 (-0.48)	100.04 (-0.28)	16.21 (-0.07)	103.74 (-0.17)	409.60 (-1.23)	363.31 (-1.79)	235.08 (-0.88)
QZP + 1d - DKH <sup>l</sup>	303.22 (-0.67)	224.39 (-0.46)	172.03 (-0.39)	101.88 (-0.25)	16.39 (0.01)	109.73 (-0.27)	419.10 (-1.32)	375.35 (-1.69)	245.69 (-0.81)
5ZP + 1d - DKH <sup>l</sup>	303.94 (-0.67)	225.36 (-0.46)	172.87 (-0.40)	102.40 (-0.26)	16.41 (0.01)	112.53 (-0.27)	421.73 (-1.33)	379.40 (-1.69)	250.02 (-0.83)
6ZP + 1d - DKH <sup>l</sup>	304.36 (-0.67)	225.87 (-0.45)	173.25 (-0.40)	102.65 (-0.25)	-	113.18 (-0.28)	423.13 (-1.35)	381.36 (-1.70)	251.99 (-0.85)
CBS(5,6 + 1d - DKH) <sup>m</sup>	304.94	226.57	173.77	102.99	16.43	114.09	425.05	384.05	254.70
CBS(5,6 + 1d - DKH) <sup>n</sup>	(-0.67)	(-0.44)	(-0.40)	(-0.24)	(0.01)	(-0.28)	(-1.38)	(-1.71)	(-0.88)
ΔE <sub>CV</sub> <sup>j</sup>	-0.2	0.4	0.4	0.3	0.3	0.8	-0.38	0.31	1.0
ΔE <sub>SO</sub> <sup>k</sup>	-0.4	0.0	-0.6	-0.8	0.0	0.0	-1.36	-3.79	-1.0
6ZP + 1d - DKH + ΔE <sub>CV</sub> + ΔE <sub>SO</sub>	303.76	226.27	173.05	102.15	16.71	113.98	421.39	377.88	251.99
CBS(5,6 + 1d - DKH) + ΔE <sub>CV</sub> + ΔE <sub>SO</sub>	304.34	226.97	173.57	102.49	16.73	114.89	423.31	380.57	254.70
Previous theoretical values <sup>o</sup>	303.7	226.6	173.7	102.0	16.8	117.8	425.1	380.6	252.8
Experimental data <sup>p</sup>	302.6±0.5	228.6±0.4	173.1±0.2	102.2±0.5	16.8±0.3	116.1±0.5	422.7±0.6	379.3±0.3	254.0±0.2

<sup>a</sup> Results are given in kcal mol<sup>-1</sup>. The experimental geometries were used in the calculations (see text). <sup>b</sup> Present investigation, basis set generated by Canal Neto et al. [8]. <sup>c</sup> Present investigation, basis set generated by Barbieri et al. [9]. <sup>d</sup> Present investigation, basis set generated by Jorge et al. [10]. <sup>e</sup> Present investigation, basis set generated by Campos et al. [11]. <sup>f</sup> Present investigation, basis set generated by Jorge et al. [18]. <sup>g</sup> Present investigation, basis set generated by Campos et al. [11]. <sup>h</sup> Present investigation, atomization energies obtained from two-point fits (5ZP-DKH and 6ZP-DKH) to Eq. 1. <sup>i</sup> Present investigation, scalar relativistic effects obtained from two-point fits (5ZP-DKH and 6ZP-DKH) to Eq. 1. <sup>j</sup> Core/valence corrections are from [19, 20]. <sup>k</sup> The atomic spin-orbit corrections are from [43]. <sup>l</sup> Present investigation, tight d function generated in this work. <sup>m</sup> Present investigation, atomization energies obtained from two-point fits (5ZP + 1d-DKH and 6ZP + 1d-DKH) to Eq. 1. <sup>n</sup> Present investigation, scalar relativistic effects obtained from 2-point fits (5ZP + 1d-DKH and 6ZP + 1d-DKH) to Eq. 1. <sup>o</sup> CBS results from [19, 20]. <sup>p</sup> Taken from JANAF (third edition) [40] and NIST-JANAF (fourth edition) [41] tables

difference between the corresponding 5ZP+1d-DKH and 6ZP+1d-DKH results is  $\leq 1\%$ . The current study clearly indicates that the computed scalar relativistic effects can display a significant basis set dependence, and that calculations with basis sets that are smaller than quadruple zeta valence quality produce results that are unreliable. For  $\text{AlF}_3$ , larger basis sets or an extrapolation scheme is necessary to estimate the CBS limits. As the extrapolation itself is cost-free compared with the calculations needed to obtain the numbers to use, it is preferable. Equation 1, along with the 5ZP+1d-DKH and 6ZP+1d-DKH scalar relativistic effect values, was used to estimate the CBS(5,6+1d-DKH) limits (see Table 3). Unless otherwise noted, these limits will be used in the discussion that follows.

For second-row systems, scalar relativistic effects become more pronounced, particularly for multihalogenated species. Feller and Dixon have determined scalar relativistic effects of  $-1.29$  and  $-1.62$  kcal mol $^{-1}$  for  $\text{AlF}_3$  and  $\text{SiCl}_4$ , respectively [19]. The scalar effect can also be important for nonhalogenated species such as  $\text{SO}_2$ , where the effect is  $-0.9$  kcal mol $^{-1}$  [20]. However, we note that the scalar relativistic effects cited above for the atomization energies of  $\text{AlF}_3$  and  $\text{SiCl}_4$  seem to be smaller than those obtained with the 6ZP+1d-DKH set and the extrapolation scheme of Eq. 1 (see Table 3). The scalar relativistic corrections reported in [19] were based on CISD-DKH/cc-pVTZ mass-velocity/one-electron Darwin calculations. Even though the scalar relativistic effect is small, Bauschlicher [42] found that using a CISD approach introduces a sizable error. Besides, our results, which were obtained at a highly accurate level of theory, showed that a basis set of triple zeta quality can also be a source of error. On the other hand, the  $\text{SO}_2$  scalar relativistic effect reported in [20] is in good agreement with our value ( $-0.88$  kcal mol $^{-1}$ ).

Before we perform a direct comparison of the CCSD(T)/6ZP+1d-DKH atomization energies and previously published experimental [40, 41] (at 0 K) and theoretical [19, 20] values, it is necessary to add some corrections to our results. (i) The atomic spin-orbit (SO) corrections,  $\Delta E_{\text{SO}}$ , were based on the tables of Moore [43]. (ii) As our binding energies were based on frozen core CCSD(T) theory, the core/valence (CV) corrections,  $\Delta E_{\text{CV}}$ , were taken from [19, 20].

The 6ZP-DKH+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$ , CBS(5,6-DKH)+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$ , 6ZP+1d-DKH+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$ , CBS(5,6+1d-DKH)+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$ , CBS [19, 20], and experimental [40, 41] atomization energies are displayed in Table 3. The CBS results [19, 20] account for valence and core correlation effects, spin-orbit and scalar relativistic effects, as well as for the inclusion of tight d functions for second-row elements. It should be mentioned here that a three-parameter extrapolation scheme was used to estimate the CBS limits reported in [19, 20], and that these limits were based on

results from the basis sets aug-cc-pVTZ through aug-cc-pV5Z. When compared with the experimental atomization energies, it is clear that there is an improvement from 6ZP-DKH+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$  to 6ZP+1d-DKH+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$  for all of the molecules studied. Besides, the latter results are generally underestimates, with the largest error being 2.00 kcal mol $^{-1}$  for  $\text{SO}_2$ . This is indicative that a larger set is needed to achieve an accuracy of  $<1$  kcal mol $^{-1}$ . However, according to our knowledge, these are the best results obtained so far with a finite basis set. It should be mentioned here that the CBS [19] error is also large (2.4 kcal mol $^{-1}$ ) for  $\text{AlF}_3$ . Overall, the mean absolute deviations (MADs) from the experimental values of the 6ZP-DKH+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$ , 6ZP+1d-DKH+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$ , and CBS [19, 20] atomization energies are 1.57, 1.17, and 1.17 kcal mol $^{-1}$ , respectively. These deviations clearly show that on average the 6ZP+1d-DKH+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$  results can be as good as those obtained from extrapolation schemes. On the other hand, the MAD decreases to 0.89 kcal mol $^{-1}$  when the CBS(5,6+1d-DKH)+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$  approach is used. All errors are then smaller than 1.75 kcal mol $^{-1}$ . Considering the uncertainty bars for the experimental data, it can decrease to 1.25 kcal mol $^{-1}$ . The worst results were obtained with CBS(5,6-DKH)+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$  (MAD was equal to 2.07 kcal mol $^{-1}$ ). The difference between the CBS(5,6-DKH)+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$  and CBS(5,6+1d-DKH)+ $\Delta E_{\text{CV}}$ + $\Delta E_{\text{SO}}$  values was 4.88 kcal mol $^{-1}$  for  $\text{SO}_2$ . This confirms the previous finding [19, 20, 26, 27] that it is not possible to obtain reliable atomization energy extrapolations for molecules containing second-row elements without adding at least a tight d function to the basis sets [XZP-DKH (XZP, X=D, T, Q, 5, and 6) in our case].

For  $\text{SO}_2$ , nonrelativistic CCSD(T) geometric parameters and harmonic vibrational frequencies obtained with different basis sets are given in Table 4. Once again, the frozen core approach to electron correlation was used. Computational cost

**Table 4** CCSD(T) geometric parameters and harmonic vibrational frequencies of  $\text{SO}_2$  with different basis sets. The frozen core approximation was used for electron correlation

	QZP <sup>a</sup>	QZP+1d <sup>b</sup>	aug-cc-pVQZ+1d <sup>c</sup>	Expt.
$r_e$ (Å)	1.4438	1.4351	1.43716	1.43076(13) <sup>d</sup>
$\theta_e$ (°)	118.65	119.22	119.25	119.33(1) <sup>d</sup>
$\omega_1$ (cm $^{-1}$ )	1158.2319	1172.6574	1166.056	1167.91(4) <sup>e</sup>
$\omega_2$ (cm $^{-1}$ )	521.0501	524.7262	519.503	522.21(3) <sup>e</sup>
$\omega_3$ (cm $^{-1}$ )	1360.1119	1384.2727	1376.171	1381.82(2) <sup>e</sup>

<sup>a</sup> Present investigation, basis set generated by Barbieri et al. [9]

<sup>b</sup> Present investigation, tight d function generated in this work

<sup>c</sup> From [27]

<sup>d</sup> From [45]

<sup>e</sup> From [46]

considerations prevented us from performing geometry optimizations and then frequency calculations at the CCSD(T)/6ZP level, but they can be performed if the basis set is reduced to QZP.

When a tight d function is added to the QZP set,  $r(\text{SO})$  shortens by 0.0087 Å and  $\theta_e$  increases by 0.57°, improving the agreement with the experimental data.

It is clear that the customary [44] accuracy of better than 10 cm<sup>-1</sup> (even for second-row compounds) for the CCSD(T)/QZP procedure is most definitely not achieved here. The deviations of the computed from the observed harmonics are -9.68, -1.16, and -20.89 cm<sup>-1</sup>. Merely adding a single tight d function dramatically cuts these errors to 4.75, 2.52, and 2.45 cm<sup>-1</sup>. Therefore, the importance of inner polarization functions is clearly not limited to the atomization energy and the geometry.

When compared with the experimental values [45, 46], the QZP+1d results are as good as the corresponding values [27] computed with the larger aug-cc-pVQZ+1d set (cf. Table 4).

## Conclusions

In this work, in order to get a better description of some molecular properties, the XZP and XZP-DKH (X=D, T, Q, 5, and 6) [8–11, 18] sets for the atoms from Na to Cl were augmented with tight d functions, which were optimized in the molecular environment at the MP2 level.

The addition of high-exponent d inner polarization functions was found to be crucial to accelerating the convergence process. The XZP+1d-DKH correlated basis sets are recommended in order to obtain a reliable extrapolation to the one-particle basis set limit. Using the extrapolation scheme of Helgaker et al. [31, 32], and taking into account inner-shell correlation and atomic spin-orbit splitting, the best atomization energy estimates were obtained with the CBS(5,6+1d-DKH)+ $\Delta E_{\text{CV}}+\Delta E_{\text{SO}}$  approach. While the scalar relativistic effect is very important in absolute terms, it is still smaller than the inner polarization effect.

Some specific conclusions can be drawn from our results:

- Even though the effect of the tight d functions on the ZPVEs is small, accounting for it helps to improve the agreement between theory and experiment
- The 6ZP+1d-DKH+ $\Delta E_{\text{CV}}+\Delta E_{\text{SO}}$  atomization energies were found to be in good agreement with those obtained from extrapolation schemes, but only the CBS(5,6+1d-DKH)+ $\Delta E_{\text{CV}}+\Delta E_{\text{SO}}$  approach reduced the atomization energy errors to levels smaller than 1.25 kcal mol<sup>-1</sup>
- Finally, good agreement between the theoretical and experimental geometric parameters and the harmonic frequencies could only be achieved when a tight d function was added to the QZP set

In summary, even when using hierarchical sequences of basis sets constructed from segmented instead of general contraction schemes, the indispensability of inner polarization functions for benchmark calculations on second-row molecules has again been highlighted.

The complete set of *s*, *p*, *d*, *f*, *g*, *h*, and *i* parameters for all basis sets for H–Ar are available online at <http://www.cce.ufes.br/qcgv/pub/>.

**Acknowledgments** We acknowledge the financial support of Conselho Nacional de Desenvolvimento Científico e Tecnológico and Fundação de Amparo à Pesquisa do Espírito Santo (Brazilian Agencies).

## References

1. Raghavachari K, Trucks GW, Pople JA, Head-Gordon M (1989) A fifth-order perturbation comparison of electron correlation theories. *Chem Phys Lett* 157:479–483
2. Dunning TH Jr (1989) Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J Chem Phys* 90:1007–1023
3. Wilson AK, van Mourik T, Dunning TH Jr (1996) Gaussian basis sets for use in correlated molecular calculations. VI. Sextuple zeta correlation consistent basis sets for boron through neon. *J Mol Struct (THEOCHEM)* 388:339–349
4. Van Mourik T, Dunning TH Jr (2000) Gaussian basis sets for use in correlated molecular calculations. VIII. Standard and augmented sextuple zeta correlation consistent basis sets for aluminum through argon. *Int J Quantum Chem* 76:205–221
5. Kendall RA, Dunning TH Jr (1992) Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J Chem Phys* 96:6796–6806
6. Jensen F (2001) Polarization consistent basis sets: Principles. *J Chem Phys* 115:9113–9125
7. Jensen F (2002) Polarization consistent basis sets. II. Estimating the Kohn–Sham basis set limit. *J Chem Phys* 116:7372–7379
8. Canal Neto A, Muniz EP, Centoducatte R, Jorge FE (2005) Gaussian basis sets for correlated wave functions. Hydrogen, helium, first- and second-row atoms. *J Mol Struct (THEOCHEM)* 718:219–224
9. Barbieri PL, Fantin PA, Jorge FE (2006) Gaussian basis sets of triple and quadruple zeta valence quality for correlated wave functions. *Mol Phys* 104:2945–2954
10. Jorge FE, Sagrillo PS, Oliveira AR (2006) Gaussian basis sets of 5 zeta valence quality for correlated wave functions. *Chem Phys Lett* 432:558–563
11. Campos CT, Ceolin GA, Canal Neto A, Jorge FE, Pansini FNN (2011) Gaussian basis set of sextuple zeta quality for hydrogen through argon. *Chem Phys Lett* 516:125–130
12. De Oliveira PJP, Jorge FE (2008) Basis-set convergence of nuclear magnetic shielding constants in molecular HF and MP2 calculations. *J Phys B: At Mol Opt Phys* 41:145101
13. Feller D, Peterson KA, de Jong WA, Dixon DA (2003) Performance of coupled cluster theory in thermochemical calculations of small halogenated compounds. *J Chem Phys* 118:3510–3522
14. Douglas M, Kroll NM (1974) Quantum electrodynamic corrections to the fine structure of helium. *Ann Phys* 82:89–155
15. Hess BA (1985) Applicability of the no-pair equation with free-particle projection operators to atomic and molecular structure calculations. *Phys Rev A* 32:756–763

16. Hess BA (1986) Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators. *Phys Rev A* 33:3742–3748
17. De Jong WA, Harrison RJ, Dixon DA (2001) Parallel Douglas–Kroll energy and gradients in NWChem: estimating scalar relativistic effects using Douglas–Kroll contracted basis sets. *J Chem Phys* 114:48–53
18. Jorge FE, Canal Neto A, Camiletti GG, Machado SF (2009) Contracted Gaussian basis sets for Douglas–Kroll–Hess calculations: estimating scalar relativistic effects of some atomic and molecular properties. *J Chem Phys* 130:064108
19. Feller D, Dixon DA (2001) Extended benchmark studies of coupled cluster theory through triple excitations. *J Chem Phys* 115:3484–3496
20. Feller D, Peterson KA (1999) Re-examination of atomization energies for the Gaussian-2 set of molecules. *J Chem Phys* 110:8384–8396
21. Kupka T (2009) Complete basis set prediction of methanol isotropic nuclear magnetic shieldings and indirect nuclear spin–spin coupling constants (SSCC) using polarization-consistent and XZP basis sets and B3LYP and BHandH density functionals. *Magn Reson Chem* 47:674–683
22. Buczek A, Kupka T, Broda MA (2011) Extrapolation of water and formaldehyde harmonic and anharmonic frequencies to the B3LYP/CBS limit using polarization consistent basis sets. *J Mol Model* 17:2029–2040
23. Buczek A, Kupka T, Broda MA (2011) Estimation of formamide harmonic and anharmonic modes in the Kohn–Sham limit using the polarization consistent basis sets. *J Mol Model* 17:2265–2274
24. Özpınar GA, Peukert W, Clark T (2010) An improved generalized AMBER force field (GAFF) for urea. *J Mol Model* 16:1427–1440
25. Özpınar GA, Kaufmann DE, Clark T (2011) Formation of the Vilsmeier–Haack complex: the performance of different levels of theory. *J Mol Model* 17:3209–3217
26. Bauschlicher CW Jr, Partridge H (1995) The sensitivity of B3LYP atomization energies to the basis set and a comparison of basis set requirements for CCSD(T) and B3LYP. *Chem Phys Lett* 240:533–540
27. Martin JML (1998) Basis set convergence study of the atomization energy, geometry, and anharmonic force field of SO<sub>2</sub>: the importance of inner polarization functions. *J Chem Phys* 108:2791–2800
28. Bauschlicher CW Jr, Ricca A (1998) Atomization energies of SO and SO<sub>2</sub>: basis set extrapolation revised. *J Phys Chem A* 102:8044–8050
29. Frisch MJ, Trucks GW, Schlegel HB et al (2009) Gaussian 09, revision A.02. Gaussian Inc., Wallingford
30. Lide DR (ed) (1994) CRC handbook of chemistry and physics. CRC, London
31. Helgaker T, Klopper W, Koch H, Noga J (1997) Basis-set convergence of correlated calculations on water. *J Chem Phys* 106:9639–9646
32. Halkier A, Helgaker T, Jorgenson P, Klopper W, Koch H, Olsen J, Wilson AK (1998) Basis-set convergence in correlated calculations on Ne, N<sub>2</sub>, and H<sub>2</sub>O. *Chem Phys Lett* 286:243–252
33. Hehre WJ, Radom L, Schleyer PvR, Pople JA (1986) *Ab initio* molecular orbital theory. Wiley, New York
34. Scott AP, Radom L (1996) Harmonic vibrational frequencies: an evaluation of Hartree–Fock, Møller–Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *J Phys Chem* 100:16502–16513
35. Merrick JP, Moran D, Radom L (2007) An evaluation of harmonic vibrational frequency scale factors. *J Phys Chem A* 111:11683–11700
36. Andrade SG, Gonçalves LCS, Jorge FE (2008) Scaling factors for fundamental vibrational frequencies and zero-point energies obtained from HF, MP2, and DFT/DZP and TZP harmonic frequencies. *J Mol Struct (THEOCHEM)* 864:20–25
37. Jackson VE, Craciun R, Dixon DA, Peterson KA, de Jong WA (2008) Prediction of vibrational frequencies of UO<sub>2</sub><sup>2+</sup> at the CCSD (T) level. *J Phys Chem A* 112:4095–4099
38. Barone V (2005) Anharmonic vibrational properties by a fully automated second-order perturbative approach. *J Chem Phys* 122:014108–014110
39. Cane E, Trombetti A (2009) The anharmonic force field of 1,3-cyclopentadienes. *Phys Chem Chem Phys* 11:2428–2432
40. Chase MW Jr, Davies AC, Downey JR Jr, Frurip DJ, McDonald RA, Syverud AN (1985) JANAF thermochemical tables, 3rd edn. *J Phys Chem Ref Data* 14 (Suppl 1)
41. Chase MW Jr (1998) NIST-JANAF tables, 4th edn. *J Phys Chem Ref Data* 9:1–1951
42. Bauschlicher CW Jr (2000) The scalar relativistic contribution to the atomization energies of CF, CF<sub>4</sub>, and SiF<sub>4</sub>. *J Phys Chem A* 104:2281–2283
43. Moore CE (1971) Atomic energy levels (NSRDS-NBS 35). Office of Standard Reference Data, National Bureau of Standards, Washington
44. Lee TJ, Scuseria GE (1995) In: Langhoff SR (ed) *Quantum mechanical electronic structure calculations with chemical accuracy*. Kluwer, Dordrecht, The Netherlands
45. Saito S (1969) Microwave spectrum of sulfur dioxide in doubly excited vibrational states and determination of the  $\gamma$  constants. *J Mol Spectrosc* 30:1–16
46. Lafferty WJ, Pine AS, Flaud J-M, Camy-Peyret C (1993) The 2 $\nu_3$  band of <sup>32</sup>S<sup>16</sup>O<sub>2</sub>: line positions and intensities. *J Mol Spectrosc* 157:499–511