JCTC Journal of Chemical Theory and Computation

Basis Set Requirements for Sulfur Compounds in Density Functional Theory: a Comparison between Correlation-Consistent, Polarized-Consistent, and Pople-Type Basis Sets

Pablo A. Denis*

DEQUIFIM, Facultad de Química, UDELAR, CC 1157, 11800 Montevideo, Uruguay

Abstract: We have investigated the SX (X = first- or second-row atom), SO_2, and SO_3 molecules employing the correlation-consistent (cc), the recently developed polarization-consistent (pc), and three Pople-type basis sets, in conjunction with the B3LYP functional. The results confirmed that the aug-pc basis sets represent a great contribution in terms of cost−benefits. In the case of the B3LYP functional, when employing the aug-pc-3 and aug-pc-4 basis sets, it is possible to obtain results that are of aug-cc-pV(5+d)Z and aug-cc-pV(6+d)Z quality, respectively, at a much lower cost. The estimations obtained employing smaller members of the family are of nearly double-ζ quality and do not provide reliable results. There is no basis set of quadruple-ζ quality among the polarized-consistent basis sets, although in terms of composition, the aug-pc-3 basis set is a QZ basis set. A precise estimation of the Kohn−Sham complete basis set (CBS) limit with the aug-pc-X basis sets is too difficult for the B3LYP functional because the ∞(aug-pc-4, aug-pc-3, aug-pc-2) extrapolation gives the same results as those of the aug-pc-4 basis set. This is in contrast with the results observed for ab initio methodologies for which the largest basis sets provided the best estimation of the CBS limit. In our opinion, the closest results to the B3LYP/CBS limit are expected to be those obtained with a two-point extrapolation employing the aug-cc-pV(X+d)Z (X = 5, 6) basis sets. The results obtained with this extrapolation are very close to those predicted by the ∞(aug-pc-3, aug-pc-2, aug-pc-1) extrapolation, and that provides a cheaper but more inaccurate alternative to estimate the CBS limit. Minor problems were found for the aug-pc-X basis sets and the B3LYP functional for molecules in which sulfur is bound to a very electronegative element, such as SO, SF, SO_2, and SO_3. For these molecules, the cc basis sets were demonstrated to be more useful. The importance of tight d functions was observed; the total atomization energy of SO_2 employing the cc-pV6Z basis set is 1 kcal/mol lower than that obtained with the aug-cc-pV(6+d)Z basis set.

Introduction

During the past decade, the computational chemistry community experienced significant changes due to the development of reliable density functionals[1]. Several theoretical groups left the traditional wave function methodologies and started to use density functional theory, in particular, the hybrid methods such as B3LYP[1]. A proof of that is that the number of articles published employing this functional, at present time, is higher than 10 000, and that number is seriously being increased every year. However, density functional theory presents a major drawback. We know that the correct functional exists, but we do not know what the correct functional is. For that reason, most of the works that employ the DFT functionals focus on resolving chemical

* Author e-mail: pablod@fq.edu.uy.

Table 1. Distances Obtained for SX, X = First-Row Atom, at the B3LYP Level of Theory Employing Different Basis Sets (Results in Å)

	number of basis functions	SB	SC	SN	SO	SF
6-31+G(d)	42	1.6216	1.5439	1.5114	1.5158	1.6411
6-311+G(3df,2p)	86	1.6101	1.5319	1.4903	1.4884	1.6137
G3Large	108	1.6093	1.5312	1.4889	1.4858	1.6084
aug-cc-pV(D+d)Z	55	1.6218	1.5434	1.5032	1.5080	1.6389
aug-cc-pV(T+d)Z	101	1.6099	1.5320	1.4909	1.4889	1.6106
aug-cc-pV(Q+d)Z	169	1.6086	1.5309	1.4884	1.4861	1.6085
aug-cc-pV(5+d)Z	263	1.6086	1.5307	1.4880	1.4853	1.6072
aug-cc-pV(6+d)Z	387	1.6086	1.5306	1.4876	1.4849	1.6069
aug-pc-0	30		1.6045	1.6113	1.6328	1.7412
aug-pc-1	50		1.5435	1.5107	1.5192	1.6516
aug-pc-2	96		1.5330	1.4930	1.4921	1.6130
aug-pc-3	178		1.5308	1.4880	1.4853	1.6073
aug-pc-4	286		1.5306	1.4877	1.4849	1.6069
∞(aug-pc-4, aug-pc-3)			1.5304	1.4874	1.4845	1.6064
∞(aug-pc-3, aug-pc-2)			1.5292	1.4844	1.4803	1.6031
∞(aug-pc-2, aug-pc-1)			1.5286	1.4855	1.4807	1.5968
∞(aug-pc-4, aug-pc-3, aug-pc-2)			1.5306	1.4877	1.4849	1.6069
∞(aug-pc-3, aug-pc-2, aug-pc-1)			1.5302	1.4860	1.4833	1.6063
∞(aug-pc-2, aug-pc-1, aug-pc-0)			1.5307	1.4890	1.4835	1.5847
experimental		1.6092[a]	1.5439[a]	1.4940[a]	1.48109[a]	1.5962[b]

[a] From ref 28. [b] From ref 29.

problems and not on the assessment of the quality of the method developed. A precise evaluation of the available DFT methods is essential for two reasons. On one hand, it helps in the development of better methodologies. On the other hand, the calibration studies permit the study of large molecular systems with smaller basis sets.

One of the first investigations of the basis set dependency in DFT was performed by Bauschlicher and Partridge in 1995[2]. They employed Pople-type and correlation-consistent basis sets[3] to estimate the total atomization energies (TAEs) of the molecules present in the G2 test set. The results showed that the TAEs determined by employing the 6-311+-(3df,2p) basis set were better than those obtained with the aug-cc-pVTZ basis set. However, for molecules that involved bondings with second-row atoms, the results were not as good as those observed for the first row. Two years later, Martell and co-workers[4] performed another study with the aim of finding the optimal combination of gradient-corrected density functionals and basis sets to study some important reactions in atmospheric chemistry. These reactions included molecules containing only the p-block elements but contained multiple bonds such as S=O. The basis sets employed for that purpose were the 6-31G(d,p), 6-311G(d,p), cc-pVDZ, and cc-pVTZ basis sets. They concluded that the latter basis set provided the best results. Some years later, Raymond and Wheeler[5] studied the compatibility of correlation-consistent basis sets and the B3LYP functional. Their results demonstrated that the minimum basis set to obtain reasonable results was the aug-cc-pVTZ and that calculations with the aug-cc-pVQZ basis set were mandatory to extrapolate to the Kohn−Sham (KS) limit. In 2000, we studied key intermediates in the oxidation of SH$_2$, namely, SO and HSO[6]. We found that, even at the B3LYP/cc-pV5Z level of theory, the results were not converged (within 2 kcal/mol) for atomization energies and bond distances (within 0.001 Å). Continuing our studies,

we published two more articles[7,8] in which we reported B3LYP calculations for more sulfur-containing molecules but increasing the size of basis set up to cc-pV6Z. The molecules studied were HSO$_2$, SO, SO$_2$, SO$_3$,[7] and the SX molecules where X = a first- or second-row atom.[8] As expected, the results confirmed the conclusions reached for HSO. The effect of diffuse functions in the estimation of relative energies employing DFT was investigated by Lynch and co-workers.[9] The authors found that the addition of diffuse functions to a double-ζ basis set was more important than increasing it to a triple-ζ basis set. Very recently, Su and co-workers[10] compared the energetics of large water clusters with basis sets up to aug-cc-pVTZ and Pople-type basis sets, namely, 6-31G** and 6-311++G**. They showed that it is possible to use the X3LYP functional and smaller basis sets while preserving accuracy, in contrast to ab initio methodologies where the basis set requirements are more important. In the organic field, Gregory and Jenks[26] employed the B3LYP functional to study the relative energies of vicinal disulfoxides and other sulfinyl radical dimers. They carried out a comparative investigation employing Pople-type basis sets and correlation-consistent basis sets up to aug-cc-pVQZ. The concluded that the aug-cc-pVQZ basis set gave results that are very close to the KS limit. Very recently, Wang and Wilson[11,22] investigated the behavior of some density functionals and correlation-consistent basis sets for a set of small first-row close-shell molecules. Wang and Wilson showed that there is a smooth convergence toward the Kohn−Sham limit for the hybrid functionals whereas the nonlocal functionals are relatively insensitive to basis set choice.

During the past five years, Jensen et al.[12−15] developed the polarized-consistent basis sets, which are specifically fitted for DFT calculations. The basic concept employed by Jensen et al. to develop the polarized-consistent basis sets is the same as that used by Dunning to construct the correlation-

Table 2. Distances Obtained for SX, X = Second-Row Atom, at the B3LYP Level of Theory Employing Different Basis Sets (Results in Å)

	SAl	SSi	SP	S_2	SCl
6-31+G(d)	2.0560	1.9517	1.9226	1.9279	2.0258
6-311+G(3df,2p)	2.0453	1.9384	1.9019	1.9026	1.9937
G3Large	2.0434	1.9362	1.9004	1.9004	1.9944
aug-cc-pV(D+d)Z	2.0557	1.9485	1.9137	1.9137	2.0076
aug-cc-pV(T+d)Z	2.0455	1.9384	1.9023	1.9023	1.9950
aug-cc-pV(Q+d)Z	2.0421	1.9356	1.9004	1.9003	1.9920
aug-cc-pV(5+d)Z	2.0430	1.9364	1.9008	1.9005	1.9918
aug-cc-pV(6+d)Z	2.0426	1.9362	1.9005	1.9001	1.9913
aug-pc-0		2.0574	2.0368	2.0985	2.2186
aug-pc-1		1.9495	1.9199	1.9267	2.0271
aug-pc-2		1.9386	1.9041	1.9052	1.9981
aug-pc-3		1.9365	1.9008	1.9005	1.9916
aug-pc-4		1.9364	1.9006	1.9001	1.9914
∞(aug-pc-4, aug-pc-3)		1.9363	1.9004	1.8897	1.9912
∞(aug-pc-3, aug-pc-2)		1.9350	1.8984	1.8971	1.9869
∞(aug-pc-2, aug-pc-1)		1.9340	1.8975	1.8962	1.9859
∞(aug-pc-4, aug-pc-3, aug-pc-2)		1.9364	1.9006	1.9001	1.9914
∞(aug-pc-3, aug-pc-2, aug-pc-1)		1.9360	1.8999	1.8992	1.9897
∞(aug-pc-2, aug-pc-1, aug-pc-0)		1.9375	1.9016	1.9019	1.9926
experimental	2.029[a]	1.9293[a]	1.9009,[a] 1.899,[b] 1.897 404[c]	1.8892[a]	1.974 65[d]

[a] From ref 28. [b] From ref 30. [c] From ref 31. [d] From ref 32.

consistent basis sets: functions that contribute similar amounts of the correlation energy are included at the same stage. However, since the convergence of HF and DFT energies is faster than that of correlation energies, the basis set optimized for describing electron correlation may not be the set of optimum exponents for describing DFT energies. To develop the aug-pc-X basis sets, Jensen et al. employed a small set of homonuclear molecules, since it is not possible to employ atoms for analyzing the importance of polarization functions. It is expected that the use of these basis set accelerates the convergence toward the KS limit. Despite the large number of articles published dealing with the basis set requirements in density functional theory, there is no investigation for second-row compounds that compares aug-pc-X and aug-cc-pV(X+d)Z basis sets employing the larger members of the families. As expressed above, the study of the basis set dependence in sulfur compounds is important because it helps in the investigation of the oxidation of sulfur in the atmosphere[4,6−7,27] and organic reactions,[23−26,37] one of our long-term research areas. In the present article, we have studied the convergence toward the KS limit of important sulfur-containing molecules, namely, the SX diatomics (X = a first- or second-row atom), SO_2, and SO_3, employing the B3LYP functional and three different families of basis sets: the augmented and nonaugmented correlation-consistent basis sets, including tight d functions,[16] aug-cc-pV(X+d)Z (X = 2, 3, 4, 5, 6); the Pople-type basis sets[17] 6-31+G(d), 6-311+G(3df), and G3Large; and the augmented polarized-consistent basis sets aug-pc-X recently developed by Jensen.[12−15]

Theoretical Methods
The hybrid DFT method B3LYP[1,18] was employed. The open-shell species have been treated employing the unre-

stricted formulation UB3LYP. Three different families of basis sets were selected to perform the comparison, (a) the aug-cc-pV(X+d)Z and cc-pV(X+d)Z correlation-consistent basis sets[16] (X = D, T, Q, 5, 6) developed by Dunning et al., (b) the augmented polarized-consistent aug-pc-X basis sets recently published by Jensen et al.,[12−15] and (c) the Pople-type basis sets 6-31+G*, 6-311+G(3df,2p), and G3Large.[17] The G3Large basis set consists of 6s 5p 4d 3f plus one diffuse sp function. Extrapolation to the Kohn−Sham limit was performed employing the two-parameter equation suggested by Halkier et al.,[21] $E_X = E_\infty + AX^{-3}$, and the three-parameter exponential extrapolation suggested by Feller,[38] $E_X = E_\infty + A\,e^{-BX}$. All calculations were carried out with Gaussian 03[19] employing the very tight optimization criteria and the ultrafine grid. The use of the default grid in Gaussian 03 instead of the ultrafine changes the geometries only by 0.0001−0.0002 Å. The basis sets were obtained from the basis set database of the Pacific Northwest National Laboratory[20] except for the aug-pc-X (X = 0, 1, 2, 3, 4) basis sets[12−15].

Results and Discussion
Bond Distances. We present the bond distances obtained for the SX, SO_2, and SO_3 molecules in Tables 1−3. A careful inspection of these tables shows that we have almost reached convergence with respect to the extension of the basis set. The aug-cc-pV(6+d)Z and aug-pc-4 results are essentially the same with only small differences of 0.0001 Å for SN, SP, and SCl and 0.0002 Å for SSi. In those cases, the aug-cc-pV(6+d)Z basis set gives smaller bond distances. Quite remarkable is the fact that the bond distances obtained with the aug-pc-3 basis set are the same as those predicted with the aug-cc-pV(5+d)Z basis set, except for three molecules, SF, SSi, and SCl. For these molecules the results employing

Table 3. Geometrical Parameters Obtained for SO_2 and SO_3 at the B3LYP Level of Theory Employing Different Basis Sets (Results in Å)

| | B3LYP | | | |
| | SO | SO_2 | | SO_3 |
	r_{S-O}	r_{S-O}	O–S–O	r_{S-O}
6-31+G(d)	1.5158	1.4655	118.65	1.4543
6-311+G(3df,2p)	1.4884	1.4367	119.22	1.4253
G3Large	1.4858	1.4346	119.23	1.4238
aug-cc-pV(D+d)Z	1.5080	1.4566	119.08	1.4440
aug-cc-pV(T+d)Z	1.4889	1.4381	119.14	1.4271
aug-cc-pV(Q+d)Z	1.4861	1.4350	119.14	1.4241
aug-cc-pV(5+d)Z	1.4853	1.4341	119.21	1.4234
aug-cc-pV(6+d)Z	1.4849	1.4336	119.23	1.4230
cc-pV(D+d)Z	1.5091	1.4525	119.31	1.4404
cc-pV(T+d)Z	1.4894	1.4370	119.25	1.4262
cc-pV(Q+d)Z	1.4864	1.4348	119.17	1.4240
cc-pV(5+d)Z	1.4854	1.4341	119.20	1.4234
cc-pV(6+d)Z	1.4849	1.4336	119.23	1.4230
aug-pc-0	1.6328	1.6051	113.64	1.6006
aug-pc-1	1.5192	1.4731	118.06	1.4608
aug-pc-2	1.4921	1.4412	118.79	1.4301
aug-pc-3	1.4853	1.4341	119.23	1.4235
aug-pc-4	1.4849	1.4336	119.24	1.4230
∞(aug-pc-4, aug-pc-3)	1.4845	1.4331	119.25	1.4224
∞(aug-pc-3, aug-pc-2)	1.4803	1.4289	119.55	1.4186
∞(aug-pc-2, aug-pc-1)	1.4807	1.4278	119.10	1.4172
∞(aug-pc-4, aug-pc-3, aug-pc-2)	1.4849	1.4336	119.24	1.4230
∞(aug-pc-3, aug-pc-2, aug-pc-1)	1.4833	1.4320	119.92	1.4216
∞(aug-pc-2, aug-pc-1, aug-pc-0)	1.4835	1.4334	118.94	1.4212
experimental	1.481 09[a]	1.4308[a]	119.3[a]	1.4173[b]

[a] From ref 33. [b] From ref 34.

the auc-cc-pV(5+d)Z basis set are shorter by 0.0001, 0.0001, and 0.0002 Å, respectively. This behavior of the aug-pc-3 basis set represents a major advantage of the polarization-consistent basis sets, since the aug-pc-3 set has 89 basis functions for sulfur and the aug-cc-pV(5+d)Z has 136. The results obtained with the aug-pc-2 basis set are of lower quality than those obtained with the aug-cc-pV(T+d)Z basis set, and that is not an advantage but, rather, a disadvantage, since the increment of the cardinal number of the basis set in one changes the results from less than triple-ζ quality to quintuple-ζ quality.

The different behavior of the correlation- and polarization-consistent basis sets is clearly seen in Figure 1 where we plot the r_{S-O} of SO_2 against the cardinal number for both families of basis sets, the correlation- and polarization-consistent basis sets. The slope of the curve fitted with the results obtained with the aug-pc-X family is much higher than that obtained with the correlation-consistent basis sets, although both converge toward the same limit. The higher slope in the curve of the aug-pc-X family makes more difficult the estimation of the Kohn–Sham limit. This is in contrast with the results obtained with the correlation-consistent basis set for which the slope is smaller, and the betterment of the results with respect to the cardinal number of basis sets is less abrupt. However, there is a problem in the convergence of the correlation-consistent basis sets, already noted by several authors.[5–7,11,22] The convergence with the correlation-consistent family is not always well-
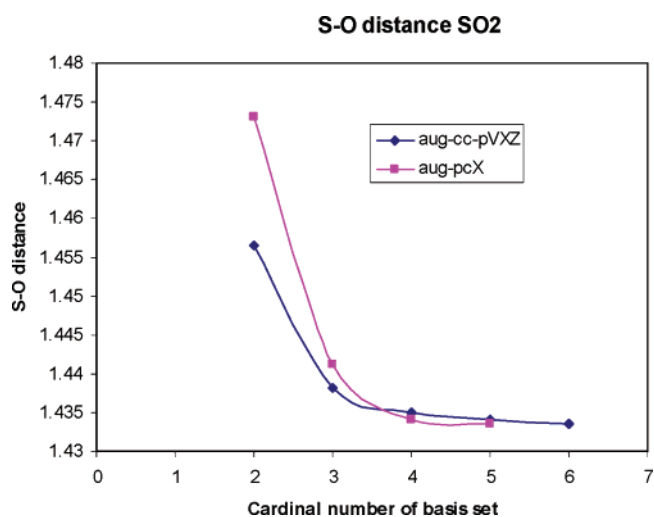
Figure 1. Convergence of the S–O distance for SO_2 with the B3LYP method and the aug-cc-pV(X+d)Z and aug-pc-X basis sets.

behaved. This can be illustrated by comparing the results obtained for the bond distances of the SX molecules where X is a first-row atom and those where X is a second-row atom (Tables 1 and 2). The convergence with respect to the basis set is good for the first row, but it is not good for the second row. For the latter, there is a problem with the aug-cc-pV(Q+Z) basis set. The predictions of the latter basis set are too short for SSi, SP, and S_2. Therefore, extrapolation

Table 4. Estimated TAEs for SX, X = First-Row Atom, at the B3LYP Level of Theory Employing Different Basis Sets (Results in kcal/mol)

	SB	SC	SN	SO	SF
6-31+G(d)	129.54	161.20	106.43	116.74	79.53
6-311+G(3df,2p)	133.08	166.47	115.11	126.41	84.91
G3Large	132.87	166.29	115.17	126.70	84.79
aug-cc-pV(D+d)Z	130.88	164.25	111.02	121.43	82.42
aug-cc-pV(T+d)Z	132.67	166.02	114.62	126.30	85.40
aug-cc-pV(Q+d)Z	133.13	166.53	115.49	126.76	85.94
aug-cc-pV(5+d)Z	132.97	166.50	115.60	127.42	86.00
aug-cc-pV(6+d)Z	133.02	166.56	115.73	127.55	86.34
∞(6, 5)	133.09	166.64	115.91	127.73	86.81
aug-pc-0		132.62	73.03	86.69	65.02
aug-pc-1		162.12	108.22	118.13	80.53
aug-pc-2		165.79	114.00	125.52	85.27
aug-pc-3		166.48	115.51	127.20	86.01
aug-pc-4		166.50	115.63	127.29	86.06
∞(aug-pc-4, aug-pc-3)		166.52	115.76	127.38	86.11
∞(aug-pc-3, aug-pc-2)		166.98	116.61	128.43	86.55
∞(aug-pc-2, aug-pc-1)		167.34	116.43	128.63	87.27
∞(aug-pc-4, aug-pc-3, aug-pc-2)		166.50	115.64	127.30	86.06
∞(aug-pc-3, aug-pc-2, aug-pc-1)		166.64	116.05	127.69	86.15
∞(aug-pc-2, aug-pc-1, aug-pc-0)		166.33	115.14	127.89	87.41
reference value[a]	134.43	171.06	113.39	124.89	84.79

[a] From ref 8. See text for details.

cannot be accomplished employing the aug-cc-pV(Q+d)Z basis set.

We have investigated two different extrapolation schemes, the two-point extrapolation and the three-point extrapolation. It is worth noticing that good results are not expected from the two-point extrapolation, since this formula is expected to be well-behaved only for correlation energies. A three-point extrapolation ∞(aug-pc-4, aug-pc-3, aug-pc-2) does not improve the results obtained with the aug-pc-4 basis set, even in 0.0001 Å. Although we did not perform numerical calculations, it is not expected that the aug-pc4 results are the real complete basis set (CBS) limits because, for some molecules, the aug-cc-pV(6+d)Z basis set provided shorter bond distances. In our opinion, this problem of the aug-pc basis sets is due to the lack of QZ quality results. In contrast, if we employ a two-point extrapolation, ∞(aug-pc-4, aug-pc-3), the bond distances are shortened with respect to the aug-pc4 results. The other two three-point extrapolation schemes investigated, ∞(aug-pc-3, aug-pc-2, aug-pc-1) and ∞(aug-pc-2, aug-pc-1, aug-pc-0), give reasonable results. In particular, the ∞(aug-pc-2, aug-pc-1, aug-pc-0) results are very impressive if we consider, first, that the extrapolated results are very close to those obtained with the aug-pc4 basis set and, second, the basis set results are of less than triple-ζ quality.

Finally, we have performed some comparative calculations with the Pople-type basis sets 6-31+G(d), 6-311+G(3df), and G3Large. The results obtained with these sets are quite good if we consider that they have much fewer basis functions. For example, for the largest basis set, G3large, the results are very close to those obtained with the aug-pc-3 basis set. It is important to remark that the G3large basis set has 66 basis functions for sulfur, only 16 more that the aug-pc-2 basis set and 20 less than the aug-pc-3 basis

set. The 6-31+G(d) basis set predicted bond distances that are longer than those obtained with the aug-cc-pV(D+d)Z basis set and very close to those obtained with the aug-pc-1 basis set. Finally, the popular 6-311+(3df,2p) basis set is of nearly aug-cc-pV(T+d)Z quality, and with respect to the polarization-consistent basis sets, it is very close to the aug-pc-2 results.

In Tables 1−3, we have included the experimental bond distances to compare them with the theoretical results obtained in the present article. For 9 of a total of 12 molecules investigated, B3LYP overestimates the bond distances with respect to the experiment. The exceptions are SB, SC, and SN. When employing the aug-cc-pV(6+d)Z basis set, the worst results are obtained for SCl, for which the bond distance predicted is 0.0166 Å larger than the experimental result.

Atomization Energies. The total atomization energies (TAEs) estimated for the molecules investigated in the present article are presented in Tables 4−6. For 8 of the 12 molecules studied in the present article, the TAEs estimated with the aug-pc-4 basis set are within 0.1 kcal/mol of the results obtained with the aug-cc-pV(6+d)Z basis set. This is a tremendous achievement if we consider that a calculation for SN employing the aug-cc-pV(6+d)Z basis set has 387 basis functions and the calculation with the aug-pc-4 basis set has 286 basis functions, 101 less. However, for the most electronegative elements oxygen and fluorine, some differences exist. This is in contrast with the results obtained for bond distances for which both basis set gave almost the same results. Employing the aug-pc-4 basis set, the TAEs obtained for SO, SF, SO_2, and SO_3 are 0.26, 0.28, 0.44, and 0.66 kcal/mol lower, respectively, than those calculated with the largest correlation-consistent basis set. This problem of the aug-pc-4 basis set is also observed when we compare the aug-

Basis Set Requirements for Sulfur Compounds

J. Chem. Theory Comput., Vol. 1, No. 5, 2005 **905**

Table 5. Estimated TAEs for SX, X = Second-Row Atom, at the B3LYP Level of Theory Employing Different Basis Sets (Results in kcal/mol)

	SAl	SSi	SP	S$_2$	SCl
6-31+G(d)	91.07	136.38	96.18	94.89	59.63
6-311+G(3df,2p)	94.56	141.93	104.72	103.05	67.02
G3Large	94.58	142.15	104.53	103.08	66.85
aug-cc-pV(D+d)Z	92.56	140.34	101.97	101.27	64.94
aug-cc-pV(T+d)Z	94.35	141.93	104.46	103.24	67.13
aug-cc-pV(Q+d)Z	94.71	142.79	104.84	103.67	67.49
aug-cc-pV(5+d)Z	94.49	142.13	104.81	103.62	67.56
aug-cc-pV(6+d)Z	94.55	142.21	104.95	103.74	67.66
∞(6, 5)	94.63	142.32	105.14	103.90	67.79
aug-pc-0		112.37	67.20	67.30	42.91
aug-pc-1		139.34	99.30	96.97	61.39
aug-pc-2		141.64	103.85	102.46	66.69
aug-pc-3		142.17	104.86	103.68	67.59
aug-pc-4		142.16	104.90	103.71	67.63
∞(aug-pc-4, aug-pc-3)			104.94	103.74	67.67
∞(aug-pc-3, aug-pc-2)		142.56	105.60	103.35	68.25
∞(aug-pc-2, aug-pc-1)		142.61	105.77	104.77	68.92
∞(aug-pc-4, aug-pc-3, aug-pc-2)			104.90	103.71	67.63
∞(aug-pc-3, aug-pc-2, aug-pc-1)		142.34	105.17	104.03	67.77
∞(aug-pc-2, aug-pc-1, aug-pc-0)		141.86	104.72	103.77	68.92
reference value[a]	97.80	146.55	103.87	103.23	68.09

[a] From ref 8. See text for details.

pc-3 and aug-cc-pV(5+d)Z basis sets. These results show that the correlation-consistent basis sets are more robust than the aug-pc-X, perhaps because of the presence of more tight d functions. Indeed, a calculation performed with the cc-pV6Z basis set yields TAE(SO$_2$) = 252.79 kcal/mol, 1 kcal/mol lower than that obtained with the aug-cc-pV(6+d)Z basis set, showing that tight d functions are also important at the DFT level, as noted by Wang and Wilson for the cc-pV5Z basis set.[36] As observed for the bond distances, the aug-pc-2 basis set gives results that are of a lesser quality than those predicted with the aug-cc-pV(T+d)Z basis set.

In Figure 2, we plot the TAE of SO$_3$ against the cardinal number of the basis set. The whole picture is similar to that discussed in bond distances. Both families of basis sets converge to the same dissociation limit, but the slope of the aug-pc-X set is larger, showing a much different performance between the members of the series. Basically, we have two zones, one below triple-ζ quality and the other above quintuple-ζ quality. Thus, the quadruple-ζ member is missing. For comparative purposes, we have also extrapolated the TAE to the KS limit for both families of basis sets. It has been observed for the CCSD(T) calculations that the extrapolations employing the largest basis sets, 5Z and 6Z, provide the best estimation of the CBS limit. However, this is not true for the aug-pcX basis sets. The ∞(aug-pc-4, aug-pc-3, aug-pc2) extrapolation does not make an important improvement, and in some cases, for example, SO, SO$_2$, and SO$_3$, the results are worse than those obtained with the aug-cc-pV(6+d)Z basis set. The latter statement is also true for a two-point extrapolation employing the aug-pc-4 and aug-pc-3 basis sets. Therefore, we employed another two alternatives to estimate the CBS limit, the use of the correlation-consistent basis sets. The results obtained with the three-point extrapolation ∞(6, 5, Q) are extremely close to those

obtained with the aug-cc-pV(6+d)Z basis set, making only a little improvement in the TAE. For example, in the case of SO$_2$, the TAE is increased by only 0.2 kcal/mol with respect to the aug-cc-pV(6+d)Z results, which is quite unlikely, since the change from aug-cc-pV(5+d)Z to aug-cc-pV(6+d)Z is 0.38 kcal/mol. The ∞(6, 5) extrapolation employing the correlation-consistent basis set appears to be quite reasonable, and it is very close to the results obtained with a three-point extrapolation ∞(aug-pc-3,aug-pc-2,aug-pc1), but because of the lack of numerical DFT calculations, we cannot be 100% sure about our observations; further investigation is required in this field to elucidate how the extrapolation should be performed.

We have investigated other three-point extrapolations, ∞(aug-pc-3, aug-pc-2, aug-pc1) and ∞(aug-pc2, aug-pc-1, aug-pc-0). The latter provides very good results at a small cost and the former, as expressed above, gives results that are very close to those obtained with a two-point extrapolation ∞(6, 5). The Pople-type basis sets give very good results, providing the best cost−results benefits. For the SX molecules, the G3Large basis set is within 1 kcal/mol with respect to the results obtained with the aug-cc-pV(6+d)Z basis set. However, it is important to remark that some unexpected deviations exist. For example, in the case of SF, the TAE estimated with the G3large basis set is 1.55 kcal/mol lower, whereas the smaller 6-311+G(3df,2p) basis set performs better. This result makes the predictions employing the Pople-type basis sets quite unsafe because they do not have uniform behavior.

Finally, it is important to compare the TAEs estimated at the DFT level with the correct results. In a recent article,[8] we investigated the thermochemistry of the SX diatomics at the CCSD(T) level, including corrections for core−valence correlation, scalar relativistic, and spin−orbit effects. The

Table 6. Estimated TAEs for SO, SO_2, and SO_3 at the B3LYP Level of Theory Employing Different Basis Sets (Results in kcal/mol)

	B3LYP		
	SO	SO_2	SO_3
6-31+G(d)	116.74	224.16	290.23
6-311+G(3df,2p)	126.41	250.61	331.78
G3Large	126.70	251.78	333.17
aug-cc-pV(D+d)Z	121.43	236.57	310.99
aug-cc-pV(T+d)Z	126.30	250.35	331.22
aug-cc-pV(Q+d)Z	127.03	252.28	334.18
aug-cc-pV(5+d)Z	127.42	253.41	335.84
aug-cc-pV(6+d)Z	127.55	253.79	336.40
∞(6, 5)	127.73	254.31	337.17
∞(5, 4)	127.89	254.60	337.58
∞(Q, T)	127.09	253.69	336.34
∞(6, 5, Q)	127.61	253.99	336.69
∞(5, Q, T)	127.88	255.04	338.03
cc-pV(D+d)Z	118.57	231.48	307.60
cc-pV(T+d)Z	125.84	250.20	332.77
cc-pV(Q+d)Z	126.85	252.33	334.86
cc-pV(5+d)Z	127.11	252.86	335.19
cc-pV(6+d)Z	127.24	253.23	335.61
aug-pc-0	86.69	136.50	153.87
aug-pc-1	118.13	225.71	293.02
aug-pc-2	125.52	248.12	327.70
aug-pc-3	127.20	253.04	335.24
aug-pc-4	127.29	253.35	335.74
∞(aug-pc-4, aug-pc-3)	127.38	253.68	336.26
∞(aug-pc-3, aug-pc-2)	128.43	256.63	340.74
∞(aug-pc-2, aug-pc-1)	128.63	257.56	342.30
∞(aug-pc-4, aug-pc-3, aug-pc-2)	127.30	253.37	335.78
∞(aug-pc-3, aug-pc-2, aug-pc-1)	127.69	254.47	337.41
∞(aug-pc-2, aug-pc-1, aug-pc-0)	127.89	255.70	339.24
reference value	124.89[a]	258.25[b]	343.72[b]

[a] From ref 8. See text for details. [b] From ref 35.
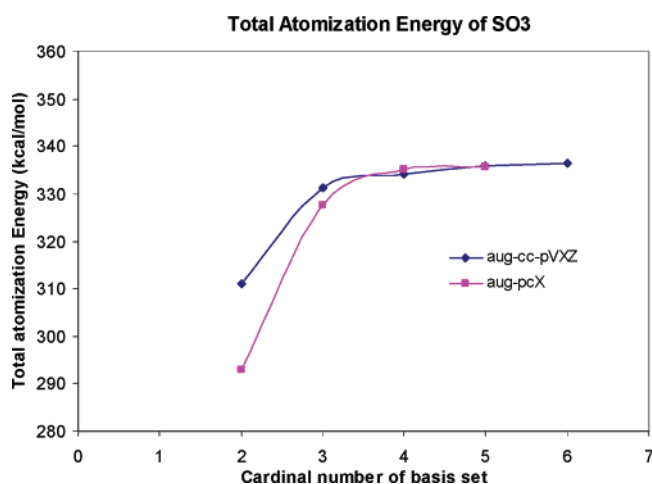
Figure 2. Convergence of the TAE for SO_3 with the B3LYP method and the aug-cc-pV(X+d)Z and aug-pc-X basis sets.

results showed that only the experimental enthalpies of formation of two molecules are known with enough accuracy to be used as a reference, namely, those of SO and SC. Errors as large as 10 kcal/mol have been found for SB, SCl, and SAl. Therefore, we will use, as a reference, the values determined at the CCSD(T) level. In the cases of SO_2 and SO_3, there are available accurate experimental TAEs.[35] These two molecules present the largest deviations. At the B3LYP/ aug-cc-pV(6+d)Z level of theory, their TAEs are 4.5 and 7.3 kcal/mol, respectively, smaller than the experimental values. This is in contrast with the results obtained for the TAE of the SX diatomics. At the same level of theory, the mean absolute deviation is 2.2 kcal/mol. Although, it is important to notice that, for three SX diatomics, we have observed deviations similar to that of SO_2, namely, for SC, SAl, and SSi. For those molecules that present larger errors, SO_2, SO_3, SC, SAl, and SSi, the deviation is by defect.

Conclusions

We have investigated the SX (X = first- or second-row atom), SO_2, and SO_3 molecules employing the correlation-consistent (cc), the recently developed polarization-consistent (aug-pc), and three Pople-type basis sets, in conjunction with the B3LYP functional. The results confirmed that the aug-pc basis sets represent a great contribution in terms of cost−benefits. In the case of the B3LYP functional, when the aug-pc-3 and aug-pc-4 basis sets are employed, it is possible to obtain results that are of aug-cc-pV(5+d)Z and aug-cc-pV-(6+d)Z quality, respectively, at a much lesser cost. The estimations obtained employing smaller members of the family are of nearly double-ζ quality and do not provide reliable results. There is no basis set of quadruple-ζ quality among the polarized-consistent basis sets, although, in terms of composition, the aug-pc-3 basis set is a QZ basis set. A precise estimation of the Kohn−Sham CBS limit with the aug-pc-X basis sets is too difficult for the B3LYP functional because the ∞(aug-pc-4, aug-pc-3, aug-pc-2) extrapolation gives the same results as the aug-pc-4 basis set. This is in contrast to the result observed for ab initio methodologies for which the largest basis set provided the best estimation of the CBS limit. In our opinion, the closest results to the B3LYP/CBS limit are expected to be those obtained with a two-point extrapolation employing the aug-cc-pV(X+d)Z (X = 5, 6) basis sets. The results obtained with this extrapolation are very close to those predicted by the ∞(aug-pc-3, aug-pc-2, aug-pc-1) extrapolation, and that provides a cheaper but more inaccurate alternative to estimate the CBS limit. Minor problems were found for the aug-pc-X basis sets and the B3LYP functional for molecules in which sulfur is bound to a very electronegative element, such as SO, SF, SO_2, and SO_3. For these molecules, the cc basis sets were demonstrated to be more useful. The importance of tight d functions was observed; the TAE of SO_2 employing the cc-pV6Z basis set is 1 kcal/mol lower than that obtained with the aug-cc-pV-(6+d)Z basis set.

References

(1) Becke, A. D. *J. Chem. Phys.* **1993**, 98, 5648.

(2) Bauschlicher, C. W.; Partridge, H. *Chem. Phys. Lett.* **1995**, *240*, 553.

Basis Set Requirements for Sulfur Compounds

J. Chem. Theory Comput., Vol. 1, No. 5, 2005 **907**

(3) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(4) Martell, J. M.; Goddard, J. D.; Eriksson, L. A. *J. Phys. Chem. A* **1997**, *101*, 1927.

(5) Raymond, K. S.; Wheeler, R. A. *J. Comput. Chem.* **1999**, *20*, 107.

(6) Denis, P. A.; Ventura, O. N. *Int. J. Quantum Chem.* **2000**, 80, 439.

(7) Denis, P. A.; Ventura, O. N. *Chem. Phys. Lett.* **2001**, 344, 221.

(8) Denis, P. A. *J. Phys. Chem. A* **2004**, *108*, 11092.

(9) Lynch, B. J.; Zhao, Y.; Truhlar D. G. *J. Phys. Chem. A* **2003**, *107*, 1384.

(10) Su, J. T.; Xu, X.; Goddard, W. A. *J. Phys. Chem. A* **2004**, *108*, 10518.

(11) Wang, N. X.; Wilson, A. K. *J. Chem. Phys.* **2004**, *121*, 7632.

(12) Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113.

(13) Jensen, F. *J. Chem. Phys.* **2002**, *116*, 7372.

(14) Jensen, F. *J. Chem. Phys.* **2002**, 117, 9234.

(15) Jensen, F.; Helgaker, T. *J. Chem. Phys.* **2004**, *121*, 3463.

(16) Dunning, T. H., Jr.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244.

(17) (a) Hehre, W.; Radom, L.; Schleyer, P. V. R. *Ab initio Molecular Orbital Theory*; Wiley: New York, 1986. (b) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.

(18) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B: Condens. Matter* **1988**, B37, 785.

(19) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision B.04; Gaussian, Inc.: Wallingford, CT, 2004.

(20) Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 2/25/04, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory, which is part of the Pacific Northwest National Laboratory, P. O. Box 999, Richland, WA 99352, and is funded by the U.S. Department of Energy. The Pacific Northwest National Laboratory is a multiprogram laboratory operated by the Battelle Memorial Institue for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact Karen Schuchardt for further information.

(21) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.

(22) Wang, N. X.; Wilson A. K. *Mol. Phys.* **2005**, *103*, 145.

(23) Gomes, J. R. B.; Ribeiro da Silva, M. A. V. *J. Org. Chem.* **2004**, *108*, 11684.

(24) Gomes, J. R. B, Gomes, P. *Tetrahedron* **2005**, *61*, 2705.

(25) McCulla, R. D.; Jenks, W. S. *J. Org. Chem.* **2003**, *68*, 7871.

(26) Gregory, D. D.; Jenks, W. S. *J. Phys. Chem. A* **2003**, *107*, 3414.

(27) Denis, P. A. *Chem. Phys. Lett.* **2005**, *402*, 289.

(28) Huber, K. P.; Herzberg, G. *Constants of Diatomic Molecules*; Van Nostrand Reinhold: New York, 1979.

(29) Endo, Y.; Nagai, K.; Yamada, C.; Hirota, E. *J. Mol. Spectrosc.* **1983**, *97*, 212.

(30) Jenouvrier, A.; Pascat, B. *Can. J. Phys.* **1978**, *56*, 1088.

(31) Kawaguchi, K.; Hirota, E.; Ohishi, M.; Suzuki, H.; Takano, S.; Yamamoto, S.; Saito, S. *J. Mol. Spectrosc.* **1988**, *130*, 81.

(32) Yamada, C.; Butler, J. E.; Kawaguchi, K.; Kanamori, H.; Hirota, E. *J. Mol. Spectrosc.* **1986**, *116*, 108.

(33) Morino, Y.; Kikuchi, Y.; Saito, S.; Hirota, R. *J. Mol. Spectrosc.* **1964**, *13*, 95.

(34) Ortigoso, J.; Escribano, R.; Maki, A. G. *J. Mol. Spectrosc.* **1989**, *138*, 602.

(35) Chase, M. W.; Davies, C. A.; Downey, J. R.; Frurip, D. J.; McDonald, R. A.; Syverud, A. N. JANAF Thermochemical Tables Third Edition. *J. Phys. Chem. Ref. Data* **1985**, *14* (Suppl. 1).

(36) Wang, N. X.; Wilson, A. K. *J. Phys. Chem. A* **2003**, *107*, 6720.

(37) Denis, P. A.; Ventura, O. N.; Mai, H. T.; Nguyen, M. T. *J. Phys. Chem. A* **2004**, *108*, 5073.

(38) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104.