11198 J. Phys. Chem. A 2007, 111, 11198−11204

# Polarization Consistent Basis Sets. 4: The Elements He, Li, Be, B, Ne, Na, Mg, Al, and Ar[†]

**Frank Jensen**

*Department of Physics and Chemistry, University of Southern Denmark, DK-5230 Odense, Denmark*

*Received: December 18, 2006; In Final Form: January 29, 2007*

Polarization consistent basis sets, optimized for density functional calculations, are proposed for the elements He, Li, Be, B, Ne, Na, Mg, Al, and Ar. The basis sets for He, B, Ne, Al, and Ar are assigned based on the previously proposed basis sets for H, C−F, and Si−Ar. The basis sets for Li, Be, Na, and Mg are defined based on energetic analysis along the lines used in previous work and the performance for molecular systems. The performance for atomization energies is comparable to those for systems composed of the elements H, C−F, and Si−Ar.

## Introduction

In previous work, we have defined a family of basis sets, denoted polarization consistent,[1−3] that are suitable for approaching the Kohn−Sham basis set limit for density functional[4] calculations in a systematic fashion. The notation is pc-*n* ($n = 0-4$), where *n* indicates the level of polarization beyond the isolated atom. The previous work defined pc-*n* basis sets for the elements H, C−F,[2] and Si−Cl,[3] and in the present case we extend these basis sets to the remaining elements in the first and second rows of the periodic table.

## Computational Details

Optimization of basis set exponents has been done by a pseudo-Newton−Raphson method using analytical gradients of the energy with respect to the basis set exponents, as implemented in the Dalton[5] program, or by numerical differences of energies, calculated by the Gaussian 03 program package.[6] As in previous work, have used the BLYP (Becke gradient corrected exchange[7] and Lee−Yang−Parr gradient corrected correlation energy[8]) functional for exponent optimization and contraction. We have previously shown that the specific choice of functional has very little influence on the performance for a selection of functionals,[9] and the pc-*n* basis sets should therefore be suitable for density functional methods in general. Benchmark calculations have been performed with Gaussian 03[6] with a grid consisting of 300 radial and 974 angular points for calculating the exchange-correlation contribution. The geometries have been taken from the B3LYP/6-31G(d,p) optimized optimized values. Open-shell species have been treated within the UHF framework.

**Defining Polarization Consistent Basis Sets.** *Elements He, B, Ne, Al, and Ar.* In previous work, we have defined polarization consistent basis sets for the elements H, C−F,[2] and Si−Cl[3] based on energetic analysis of basis functions with different angular momentum. The s and p exponents are optimized at the atomic level, whereas polarization exponents are assigned based on molecular calculations. The final set of polarization exponents necessarily represents suitable average values, as each individual molecule has its own optimum set of exponents. The variation of the polarization exponents with the nuclear charge is sufficiently regular to allow a straightforward extrapolation to the B/Ne and Al/Ar elements, which together with atomic optimized s and p exponents define the pc-*n* basis sets for these elements. The benchmark calculations discussed below show that the performance is very similar to those reported earlier,[2,3] thereby justifying the procedure.

A minor irregularity occurred for the atomic Al pc-4 basis set which is 21s16p in composition. For the elements Si−Ar, the optimized exponents for 21 s functions have a 17/4 partition between the core and valence space, whereas a fully optimized basis set for Al has a 18/3 partition, and it was not possible to locate an alternative 17/4 solution. A fully optimized 22s16p basis set for Al has a 18/4 partition for the s functions, and since we felt that basis set balance between elements is important, we have used the four s valence functions from the 22s16p and optimized the remaining 17 core s functions to define the atomic pc-4 basis set for Al. This constrained optimized basis set gives an energy 2 micro-Hartree above the fully optimized 21s16p basis set. In the process of optimizing the basis sets for Al, it was discovered that the optimization of the pc-3 basis set for Si reported previously[3] had not fully converged. A reoptimization lowered the energy by 3 micro-Hartree, which is negligible in terms of energy, but placed the Si exponents in the expected region compared to those for P and Al.

The atomic s basis for helium and the number of polarization functions were assigned analogous to the pc-*n* basis sets for hydrogen. The polarization exponents for helium are not easy to define, as the polarization functions for the pc-*n* basis sets primarily serve to describe molecular bonding, and there are no known helium compounds with shared electron bonding. For neon and argon, we have relied on extrapolation from the elements with lower atomic numbers for assigning the polarization exponents, but for helium the only "extrapolation" point is hydrogen. We have instead used the ratio between the exponents for the outer s function of hydrogen and helium as a guide. This ratio follows closely the corresponding s ratio between the elements Si and Ar, and we have consequently multiplied the hydrogen polarization exponents with the ratio between the corresponding polarization functions for Si and Ar.

It should be noted that the optimum polarization exponents depend significantly on interatomic distances and are assigned based on molecular bonding. The polarization exponents for the rare gas elements (He, Ne, and Ar) are therefore "tight"

---

[†] Part of the "Thom H. Dunning. Jr., Festschrift".

10.1021/jp068677h CCC: $37.00

**TABLE 1: Systems Used as Test Set**

$Li_2$, LiH, $LiCH_3$, $Li_2C_2$, $LiNH_2$, $Li_2NH$, $Li_2O$, $Li_2O_2$, LiOH, LiF, $Li_2Si_2$, $LiPH_2$, $Li_2PH$, $Li_2S_2$, LiSH, LiCl
$Be_3$, $BeH_2$, $Be_2H_4$, $BeC_2$, $Be(CH_3)_2$, BeO, $Be(OH)_2$, $BeF_2$, $BeSi_2$, BeS, $Be(SH)_2$, $BeCl_2$
$B_2$, BH, $BH_3$, $B_2H_6$, $B(CH_3)_3$, BN, $B(NH_2)_3$, $H_2BNH_2$, $B_3N_3H_6$, BO, $B_2O_3$, $B(OH)_3$, BF, $BF_3$
$Na_2$, NaH, $NaCH_3$, $Na_2C_2$, $NaNH_2$, $Na_2NH$, $Na_2O$, $Na_2O_2$, NaOH, NaF, $NaPH_2$, $Na_2PH$, $Na_2S$, NaSH, NaCl
$Mg_3$, MgH, $MgH_2$, $Mg_2H_4$, $MgC_2$, $Mg(CH_3)_2$, MgO, $Mg(OH)_2$, $MgF_2$, MgS, $Mg(SH)_2$, $MgCl_2$, $CH_3MgCl$
$Al_2$, AlH, $AlH_3$, $Al_2H_6$, $Al_2O_3$, $AlF_3$, AlSi, $Al(SiH_3)_3$, AlP, $Al(PH_2)_3$, $H_2AlPH_2$, $Al_3P_3H_6$, AlS, $Al_2S_3$, $Al(SH)_3$, AlCl, $AlCl_3$, $Al_2Cl_3$

and will be suitable for describing the charge deformation occurring at short intermolecular distances, i.e., in the repulsive part of the interaction. Nobel gas elements have been used as probes for steric interactions, and the pc-*n* basis sets should be suitable for such applications. The regular pc-*n* basis sets are not expected to be useful for describing dispersion interactions, where augmentation with diffuse functions (aug-pc-*n* basis sets) is recommended. It should be noted that density functionals accounting for dispersion have only recently been proposed.[10]

*Elements Li and Be.* The selection of atomic s basis sets for the Li and Be elements required some considerations. The pc basis sets are designed such that there is an energetic balance between the different types of functions for a given atom. In addition, we feel that it is important that there is a balance in the basis sets across a row in the periodic table. This suggests that the number of s functions for the Li and Be elements should match the remaining first row elements and that the valence region is represented by comparable accuracy for all elements. Since we use a variational procedure for optimizing the atomic s exponents, this leads in some cases to ambiguities.

The pc-0 basis set has five s functions, and a full optimization produces exponents with a 4/1 distribution between the core and valence space for all of the first row elements. The pc-1 basis set has 7 s-functions with a 5/2 core/valence exponent distribution for the elements Be−Ne, whereas the lowest energy solution for Li has a 6/1 distribution. An alternative solution 0.7 milli-Hartree higher in energy with a 5/2 partitioning could also be optimized and was selected for consistency. The pc-2 basis has 10 s functions, and the fully optimized set of functions has exponents with a 7/3 core/valence partitioning for the elements B−Ne. Two different solutions exist for Be with 8/2 and 7/3 exponent partitionings, and we selected the latter, being 83 micro-Hartree higher in energy than the former. There are similarly two solutions for Li, but the one with a 7/3 exponent partitioning is on the verge of variational collapse, as the exponent ratio between the second and third outermost function is only 1.66. In order to ensure a consistent representation of the valence part, we have used a procedure analogous to that used for Al (vide supra) by taking the three outer valence functions from a fully optimized 12s basis (which has a 9/3 core/valence partitioning) and only optimized the innermost seven s functions. This partly optimized basis set has an atomic energy 181 micro-Hartree above the lowest energy solution.

The pc-3 basis set has 14 s functions, corresponding to a 10/4 core/valence partitioning for B−Ne, whereas the lowest energy solution for Li and Be corresponds to a 11/3 partitioning. Alternative solutions corresponding to a 10/4 partitioning also exist, but the minimum exponent ratios of 1.72 and 1.78 indicate limiting variational stability, and we opted again for taking the 4s valence part from a fully optimized 15s basis, and reoptimizing only the 10 core functions. These partly optimized 14s basis sets are 7 micro-Hartrees higher in energy than the fully optimized solutions.

The pc-4 basis set has 18 s functions for the element B−Ne, but attempts of optimizing a basis set with 18 s functions for Li and Be failed due to variational collapse. We instead chose to use the procedure proposed by Petersson et al.[11] and optimized a 19 s function basis set where the exponents were parametrized by a fourth order Legendre polynomial. Our experience is that a basis set with $M + 1$ functions defined by a fourth order Legendre polynomial typically provides results corresponding to a fully optimized basis set with $M$ functions. The exponent range for these 19s basis sets follow the regular pattern defined by the remaining first row elements.

The elements Li and Be have only occupied s orbitals, and p functions are therefore the first set of polarization exponents, in analogy with the basis sets for hydrogen. In contrast to hydrogen, however, Li and Be have core orbitals, which furthermore are sufficiently high in energy to participate in molecular bonding to some extent. The virtual p orbitals are similarly low in energy and will participate in molecular bonding. It is therefore not clear whether the optimum composition of the polarization functions will be the same as for the other elements.

We have performed analysis of the energetic importance of the polarization functions for the $Li_2$ and $Be_3$ systems, along the lines used in previous work.[2,3] The $Be_3$ results are shown in Figure 1 and indicate that the consistent levels of polarization functions are 1p, 3p1d, 6p2d1f, and 8p3d2f1g. Very similar results were obtained for the $Li_2$ case. Compared to the polarization functions for the pc basis sets for the remaining first row elements (1d, 2d1f, 4d2f1g, and 6d3f2g1h), the analysis indicates that the s-block elements should have one or two extra functions for the first polarization type. This is not unreasonable, based on the contribution of core polarization to the molecular bonding and participation of the virtual p orbitals. A straightforward combination of the s and p polarization functions would suggest a basis set composition of 5s for pc-0, 7s1p for pc-1, 10s3p1d for pc-2, etc. The energy analysis in Figure 1, however, indicates that the first p function is of similar importance as

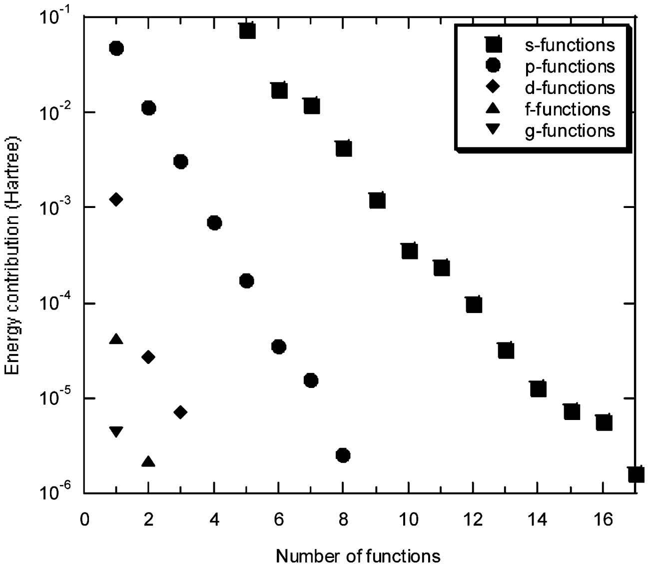

**Figure 1.** Energy contributions (Hartree) per atom of different basis functions for the $Be_3$ system.

**TABLE 2: Mean and Maximum Absolute Deviations of Atomization Energies Relative to Molecular Systems Compared to the Basis Set Limit (kJ/mol) for the Systems in Table 1**[a]

| element | Li | | Be | |
|---|---|---|---|---|
| basis | MAD | MaxAD | MAD | MaxAD |
| 5s | 102 | 400 | 275 | 521 |
| 5s1p | 50 | 157 | 64 | 167 |
| 5s2p | 41 | 210 | 42 | 113 |
| 5s3p | 77 | 276 | 51 | 150 |
| 7s1$_\mathrm{p}$ | 51 | 168 | 52 | 155 |
| 7s2p | 29 | 92 | 29 | 83 |
| 7s3p | 9 | 23 | 16 | 43 |
| 7s4p | 8 | 22 | 11 | 36 |
| 10s3p1d | 4 | 14 | 5 | 14 |
| 10s4p1d | 4 | 13 | 3 | 8 |

[a] Basis set limit results are estimated by extrapolation of the uncontracted pc-2, pc-3, and pc-4 results.

the fifth s function, and the second p function is of similar importance as the seventh s function. Based on an energy balance between the s and p polarization functions, this suggests that the pc-0 basis set should be 5s1p in composition, whereas the pc-1 should be 7s2p; that is, the pc-0 basis set should formally include polarization functions.

As the quality of the basis set ultimately is going to be judged based on performance for molecular systems, we decided to let the selection of polarization functions be guided by the performance for atomization energies relative to a molecular reference ($X_2$ and $Be_3$ molecules) for the systems shown in Table 1. The results in Table 2 show that a single p function combined with the five s functions in the pc-0 basis set reduces the error by a factor of 2−3 and produces results that are comparable to the other first row elements (see Tables 6 and 7 below). Adding a second p function gives less improvement and actually increases the MaxAD for the Li data. A third p function is clearly counter-productive and increases the error as the basis set becomes unbalanced. The pc-1 basis set has seven s functions, and the results in Table 2 indicate that it should be combined with three p functions in order to produce results of comparable quality as the other first row elements. Figure 1 indicates either a 3p1d or 4p1d polarization for the pc-2 basis set (10s), and the performance in Table 2 lead us to choose the latter.

Based on the above analysis, we thus define the pc-0 basis set as 5s1p, the pc-1 as 7s3p, the pc-2 as 10s4p1d, the pc-3 as 14s6p2d1f, and the pc-4 as 19s8p3d2f1g in composition. The polarization exponents were determined as suitable averages based on fully optimized exponents for the systems shown in Table 1.[12]

It should be noted that the pc-*n* basis sets for Li and Be are somewhat different from other popular basis sets, as these often include p-functions as part of the atomic basis (Tables 3 and 4). The Pople (STO-3G,[13] 6-31G(d,p),[14] and 6-311G(2df,2pd)[15]) and Dunning (cc-pVXZ)[16] type basis sets include a full set of valence p functions along with higher order polarization functions analogous to the other first row elements. As a result, these include polarization functions with one higher angular momentum than the pc-*n* basis sets. The SVP,[17] TVZ,[18] and

**TABLE 3: Basis Set Composition in Terms of Uncontracted Functions**

| basis | Li | Be | B/Ne | Na | Mg | Al/Ar |
|---|---|---|---|---|---|---|
| pc−0 | 5s1p | 5s1p | 5s3p | 8s5p | 8s5p | 8s6p |
| pc−1 | 7s3p | 7s3p | 7s4p1d | 11s7p | 11s7p | 11s8p1d |
| pc−2 | 10s4p1d | 10s4p1d | 10s6p2d1f | 13s9p1d | 13s9p1d | 13s10p2d1f |
| pc−3 | 14s6p2d1f | 14s6p2d1f | 14s9p4d2f1g | 17s12p2d1f | 17s12p2d1f | 17s13p4d2f1g |
| pc−4 | 19s8p3d2f1g | 19s8p3d2f1g | 18s11p6d3f2g1h | 21s15p3d2f1g | 21s15p3d2f1g | 21s16p6d3f2g1h |
| STO-3G | 6s3p | 6s3p | 6s3p | 9s6p | 9s6p | 9s6p |
| 6-31G(d,p) | 10s4p1d | 10s4p1d | 10s4p1d | 16s10p1d | 16s10p1d | 16s10p1d |
| 6-311G(2df,2pd)[a] | 11s5p2d1f | 11s5p2d1f | 11s5p2d1f | 13s9p2d1f | 13s9p2d1f | 13s9p2d1f |
| cc-pVDZ | 9s4p1d | 9s4p1d | 9s4p1d | 12s8p1d | 12s8p1d | 12s8p2d |
| cc-pVTZ | 11s5p2d1f | 11s5p2d1f | 11s5p2d1f | 16s10p2d1f | 15s10p2d1f | 15s10p3d1f |
| cc-pVQZ | 12s6p3d2f1g | 12s6p3d2f1g | 12s6p3d2f1g | 19s12p3d2f1g | 16s12p3d2f1g | 16s12p4d2f1g |
| cc-pV5Z | 14s7p4d3f2g1h | 14s7p4d3f2g1h | 14s7p4d3f2g1h | 19s12p4d3f2g1h | 20s14p4d3f2g1h | 20s14p5d3f2g1h |
| SVP | 7s1p | 7s4p | 7s4p1d | 10s6p | 10s6p | 10s7p1d |
| TZV | 11s3p | 11s4p | 11s6p2d1f | 14s8p1d | 14s8p1d | 14s9p2d1f |
| QZV | 15s6p2d1f | 15s7p2d1f | 15s7p2d1f | 20s12p3d1f | 20s12p3d1f | 20s14p4d2f1g |

[a] McLean−Chandler for Na, Mg, Al, and Ar.

**TABLE 4: Basis Set Composition in Terms of Contracted Functions**

| basis | Li | Be | B/Ne | Na | Mg | Al/Ar |
|---|---|---|---|---|---|---|
| pc−0 | 3s1p | 3s1p | 3s2p | 4s1p | 4s1p | 4s3p |
| pc−1 | 3s2p | 3s2p | 3s2p1d | 4s2p | 4s2p | 4s3p1d |
| pc−2 | 4s2p1d | 4s2p1d | 4s3p2d1f | 5s3p1d | 5s3p1d | 5s4p2d1f |
| pc−3 | 6s3p2d1f | 6s3p2d1f | 6s5p4d2fg1 | 6s4p2d1f | 6s4p2d1f | 6s5p4d2f1g |
| pc−4 | 8s4p3d2f1g | 8s4p3d2f1g | 8s7p6d3f2g1h | 7s5p3d2f1g | 7s5p3d2f1g | 7s6p6d3f2g1h |
| STO-3G | 2s1p | 2s1p | 2s1p | 3s2p | 3s2p | 3s2p |
| 6-31G(d,p) | 3s2p1d | 3s2p1d | 3s2p1d | 4s3p1d | 4s3p1d | 4s3p1d |
| 6-311G(2df,2pd)[a] | 4s3p2d1f | 4s3p2d1f | 4s3p2d1f | 6s5p2d1f | 6s5p2d1f | 6s5p2d1f |
| cc-pVDZ | 3s2p1d | 3s2p1d | 3s2p1d | 4s3p1d | 4s3p1d | 4s3p2d |
| cc-pVTZ | 4s3p2d1f | 4s3p2d1f | 4s3p2d1f | 5s4p2d1f | 5s4p2d1f | 5s4p3d1f |
| cc-pVQZ | 5s4p3d2f1g | 5s4p3d2f1g | 5s4p3d2f1g | 6s5p3d2f1g | 6s5p3d2f1g | 6s5p4d2f1g |
| cc-pV5Z | 6s5p4d3f2g1h | 6s5p4d3f2g1h | 6s5p4d3f2g1h | 7s6p4d3f2g1h | 7s6p4d3f2g1h | 7s6p5d3f2g1h |
| SVP | 3s1p | 3s2p | 3s2p1d | 4s2p | 4s2p | 4s3p1d |
| TZV | 5s2p | 5s3p | 5s3p2d1f | 5s4p1d | 5s4p1d | 5s4p2d1f |
| QZV | 6s4p2d1f | 7s4p2d1f | 7s4p3d2f1g | 9s5p3d1f | 9s5p3d1f | 9s6p4d2f1g |

[a] McLean−Chandler for Na, Mg, Al, and Ar.

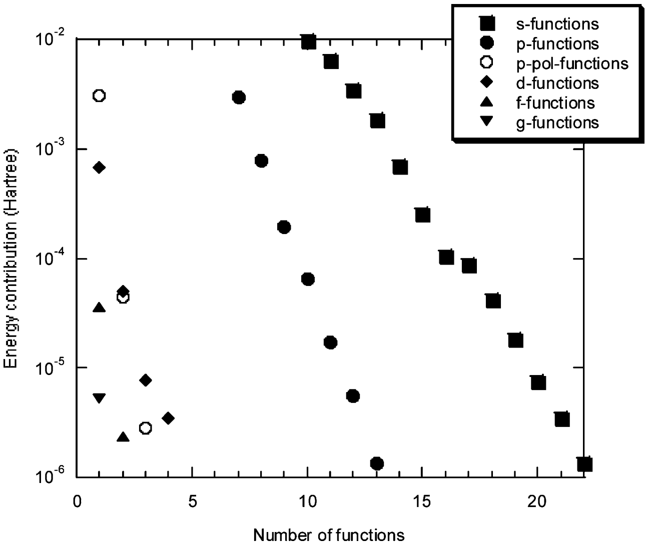

**Figure 2.** Energy contributions (Hartree) per atom of different basis functions for the $Mg_3$ system. The open circles labeled p-pol-functions are the p-type polarization functions, as opposed to the atomic p-functions labeled with solid circles.

QZV[19] basis sets developed by Ahlrichs and co-workers are more in line with the present basis sets, except that the TZP basis set does not include d functions. They are furthermore characterized by having a different number of p-type functions for Li and Be, which also differ from the remaining first row elements. Another difference is that the polarization functions for the pc-*n* basis set, especially the p functions, tend to have larger exponents than most other basis sets. This reflects the fact that the exponents in the present case have been optimized for describing molecular bonding, whereas other basis sets typically have optimized the p-type exponents for a P-type excited-state for the atom.

*Elements Na and Mg.* For Na and Mg, the atomic p basis must be smaller than for the elements Al−Ar, since there are no 3p electrons. An analysis of the energetic contributions and exponents with respect to the atomic number suggested that the pc-0 and pc-1 should have two fewer p functions, the pc-2 and pc-3 should have three fewer p functions, and the pc-4 should have four fewer p functions. The pc-0 basis set is thus defined as being 8s4p in composition, whereas the pc-1 basis set is 11s6p. The shell distribution of the s exponents for the latter is 7/2/2 for the elements Mg−Ar, whereas it is 7/3/1 for Na. No alternative solution could be found for Na, and we opted for taking the outer two s exponents from a 12s6p set of functions, which has a 7/3/2 partitioning, and reoptimize the remaining nine s exponents to provide a 7/2/2 exponent distribution for Na. This partly constrained basis set is 1.8 milli-Hartree higher in energy than the fully optimized one.

The pc-2 basis set is 13s7p in composition, while the pc-3 basis set is 17s10p. The latter has a 11/3/3 exponent distribution for the s-functions for the elements Mg−Ar, while for Na there are two solutions with the exponents distributed either 11/3/3 or 12/3/2 between the shells. The former is 28 micro-Hartree higher in energy, but was selected for consistency. For the pc-4 basis set, it was in analogy with the Al case not possible to obtain a 21s basis with a 17/4 core/valence distribution for Na and Mg. A 18/4 exponent distribution could be obtained with a 22s12p basis set, and we selected the four valence s functions from this and reoptimized only the inner 17 s functions. These partly optimized basis sets are 6 and 4 micro-Hartree higher in energy (Na and Mg, respectively) than the fully optimized solutions.

**TABLE 5: Mean and Maximum Absolute Deviations of Atomization Energies Relative to Molecular Systems Compared to the Basis Set Limit (kJ/mol) for the Systems in Table 1**[a]

| element | Na | | Mg | |
|---|---|---|---|---|
| basis | MAD | MaxAD | MAD | MaxAD |
| 8s4p | 54 | 239 | 144 | 479 |
| 8s4p1p | 56 | 201 | 42 | 123 |
| 8s4p2p | 62 | 204 | 53 | 93 |
| 8s4p3p | 64 | 206 | 54 | 93 |
| 11s6p1p | 17 | 39 | 36 | 83 |
| 11s6p2p | 12 | 32 | 25 | 67 |
| 11s6p1p1d | 10 | 31 | 17 | 53 |
| 13s7p1p1d | 7 | 18 | 8 | 21 |
| 13s7p2p1d | 4 | 12 | 5 | 17 |

[a] Basis set limit results are estimated by extrapolation of the uncontracted pc-2, pc-3 and pc-4 results.

**TABLE 6: Mean and Maximum Absolute Deviations (MAD and MaxAD) Atomization Energies Relative to Diatomic Systems Compared to the Basis Set Limit (kJ/mol)**[a]

| element | B | | Al | |
|---|---|---|---|---|
| basis | MAD | MaxAD | MAD | MaxAD |
| STO-3G | 169.5 | 626.8 | 323.4 | 853.7 |
| pc−0 | 152.3 | 533.3 | 46.5 | 165.1 |
| pc−0 (uncontracted) | 157.2 | 540.8 | 46.2 | 161.8 |
| 6-31G(d,p) | 11.9 | 39.4 | 16.7 | 92.7 |
| cc-pVDZ | 15.5 | 42.5 | 15.3 | 65.2 |
| SVP | 20.3 | 79.7 | 29.9 | 91.8 |
| pc−1 | 6.3 | 15.4 | 12.0 | 51.7 |
| pc−1 (uncontracted) | 6.8 | 25.3 | 9.7 | 48.1 |
| 6-311G(2df,2pd)[a] | 7.5 | 23.3 | 8.0 | 20.3 |
| cc-pVTZ | 2.6 | 12.9 | 2.1 | 9.8 |
| TZV | 2.3 | 7.1 | 6.7 | 34.8 |
| pc−2 | 2.7 | 8.4 | 4.7 | 17.2 |
| pc−2 (uncontracted) | 2.6 | 8.4 | 3.9 | 13.6 |
| cc-pVQZ | 2.2 | 6.4 | 2.5 | 5.7 |
| QZV | 0.7 | 1.9 | 0.6 | 2.4 |
| pc−3 | 0.5 | 1.3 | 0.3 | 1.3 |
| pc−3 (uncontracted) | 0.4 | 1.1 | 0.4 | 1.5 |
| cc-pV5Z | 0.8 | 2.2 | 0.6 | 2.1 |
| pc−4 | 0.05 | 0.12 | 0.06 | 0.16 |
| pc−4 (uncontracted) | 0.05 | 0.11 | 0.04 | 0.20 |

[a] Basis set limit is estimated by extrapolation of the uncontracted pc-2, pc-3 and pc-4 results. [b] McLean−Chandler for Al.

The elements Na and Mg have in analogy with Li and Be only occupied valence s orbitals, but for these elements, there are occupied core p orbitals as well. To the extent that the 2p basis functions can act as polarization for the 3s orbital, and for describing polarization of the 2p orbitals, d-type functions may also be considered as the first set of polarization functions. In order to probe the energetic importance of the polarization function, we have performed energetic analysis of the $Na_2$ and $Mg_3$ systems, with the results for $Mg_3$ shown in Figure 2 using the pc-4 (21s12p) atomic basis set. A single p function is the most important as it serves to polarize the 3s orbital, but the exact energy contribution of course depends on the size of the underlying p basis set. The second level of polarization is 1p1d followed by 2p2d1f. At the pc-4 level, we have decided on a 3p3d2f1g polarization space, although one could argue for including a fourth d function based ion Figure 2. The order of polarization functions is consistent with the first row s-block elements, except that the need for p-type functions is strongly reduced. This is of course due to the p functions describing the

**TABLE 7: Mean and Maximum Absolute Deviations (MAD and MaxAD) Atomization Energies Relative to Diatomic Systems Compared to the Basis Set Limit (kJ/mol)[a]**

| element | Li | | Be | | Na | | Mg | |
|---|---|---|---|---|---|---|---|---|
| basis | MAD | MaxAD | MAD | MaxAD | MAD | MaxAD | MAD | MaxAD |
| STO-3G | 102.0 | 285.1 | 99.8 | 169.6 | 297.7 | 511.9 | 224.2 | 482.2 |
| pc−0 | 50.2 | 155.9 | 62.3 | 165.5 | 49.4 | 189.0 | 38.8 | 151.5 |
| pc−0 (uncontracted) | 50.0 | 156.6 | 63.8 | 166.6 | 49.6 | 186.4 | 40.3 | 144.3 |
| SVP | 52.6 | 142.4 | 17.2 | 43.7 | 23.9 | 65.1 | 58.5 | 156.4 |
| pc−1 | 10.6 | 28.9 | 19.3 | 55.7 | 18.7 | 47.1 | 42.4 | 98.3 |
| pc−1 (uncontracted) | 8.6 | 22.5 | 16.4 | 43.2 | 16.9 | 38.6 | 35.6 | 82.8 |
| TZV | 8.6 | 19.3 | 14.6 | 38.4 | 9.6 | 28.7 | 17.9 | 56.9 |
| 6-31G(d,p) | 14.0 | 42.4 | 14.7 | 48.0 | 14.5 | 45.5 | 15.5 | 63.8 |
| cc-pVDZ | 8.8 | 36.4 | 14.2 | 45.1 | 8.6 | 37.7 | 16.0 | 63.8 |
| pc−2 | 4.7 | 15.1 | 3.2 | 7.2 | 5.0 | 13.7 | 7.2 | 23.5 |
| pc−2 (uncontracted) | 4.0 | 12.6 | 3.1 | 7.6 | 4.4 | 11.9 | 4.9 | 16.7 |
| 6-311G(2df,2pd)[a] | 4.9 | 11.0 | 5.3 | 14.4 | 4.4 | 12.1 | 4.6 | 9.3 |
| cc-pVTZ | 2.1 | 6.0 | 3.0 | 8.7 | 5.2 | 11.0 | 6.1 | 21.0 |
| QZV | 1.7 | 3.5 | 1.8 | 4.9 | 1.6 | 3.9 | 0.7 | 1.5 |
| pc−3 | 0.4 | 1.4 | 0.3 | 0.9 | 1.0 | 2.3 | 0.9 | 2.0 |
| pc−3 (uncontracted) | 0.3 | 1.4 | 0.2 | 1.0 | 0.8 | 2.1 | 0.7 | 1.7 |
| cc-pVQZ | 1.5 | 5.2 | 1.0 | 2.2 | 3.9 | 7.3 | 2.2 | 7.0 |
| pc−4 | 0.05 | 0.16 | 0.06 | 0.13 | 0.29 | 0.76 | 0.39 | 0.68 |
| pc−4 (uncontracted) | 0.04 | 0.17 | 0.03 | 0.14 | 0.11 | 0.29 | 0.10 | 0.23 |
| cc-pV5Z | 0.3 | 1.2 | 0.4 | 0.8 | 1.1 | 2.9 | 2.9 | 7.7 |

[a] Basis set limit is estimated by extrapolation of the uncontracted pc-2, pc-3 and pc-4 results. [b] McLean−Chandler for Na and Mg.

2p orbital doubles as polarization functions for the inner part of the 3s orbital.

In analogy with the Li and Be elements, we decided to guide the final selection of p-type polarization functions by the performance for molecular atomization energies, with the results shown in Table 5. The results indicate a large improvement in performance for the pc-0 basis set (8s4p) by including a single p function, especially for the Mg systems. For the pc-1 type basis set (11s6p), a second p-type function is less important than the lack of d-type functions which first enters at the pc-2 stage, and only a single p-type polarization function is therefore included for the pc-1 basis set. For the pc-2 basis set (13s7p) the analysis in Figure 2 indicates a polarization level of 1p1d, but the results in Table 5 shows that a polarization of 2p1d provided a performance more in line with the results for the remaining elements. When the atomic basis sets are combined with these choices of polarization functions, we thus define the pc-0 basis set as 8s5p, the pc-1 as 11s7p, the pc-2 as 13s9p1d, the pc-3 as 17s12p2d1f, and the pc-4 as 21s15p3d2f1g. Compared to the remaining second row elements, the basis sets for Na and Mg contain the same number of s functions and one less p function and the level of polarization functions is reduced by one (Table 3). The exponents for the polarization functions, which include the outer p-type functions as well, were selected based on explicit optimized results for the systems shown in Table 1. The exception is the three p-polarization functions for the pc-4 basis set, which were assigned as an even-tempered sequence with a ratio of 2.2, as explicit optimization lead to small exponent ratios.

## Contraction and Augmentation

The primitive set of basis functions defined in the previous section was contracted by a general contraction scheme using coefficients from atomic calculations. The contraction of the s functions follows the previous schemes,[1,3] whereas the contraction of the p functions was determined by the condition that the contraction error should be smaller than the inherent error of the uncontracted basis set relative to the basis set limit for the selection of molecular systems in Table 1. For the Li and Be elements, it was possible to substantially contract the inner p functions without losing accuracy, confirming that the p functions to a large extent serve to describe the virtual 2p orbital, rather than being polarization functions for the 2s orbital. For the Na and Mg elements, the p functions could be contracted to one less than the remaining second row elements, in agreement with expectations. The final basis set compositions are shown in Table 4.

For properties depending on the electron density far from the nuclei, for example, electric dipole moments and polarizabilities, the basis set convergence can be significantly improved by adding diffuse functions with small basis-set exponents (aug-pc-*n*). By analogy with previous work,[20] we have assigned diffuse s and p exponents ($\zeta_{\mathrm{dif}}$) from the two outer exponents of the pc-*n* basis set ($\zeta_2$ and $\zeta_1$) by the formula given in eq 1, with $K = 0.20$

$$\zeta_{\mathrm{dif}} = \frac{\zeta_1}{(\zeta_2/\zeta_1 + K)} \tag{1}$$

An exception is the exponent of the diffuse s function for the pc-0 basis set, since $\zeta_2/\zeta_1$ is very large in this case, due to the exponent gap between the core and valence orbitals. For this special case, we have chosen the ratio from the corresponding p functions or estimated it from the neighboring elements in the case of Li and Be. The exponents for the diffuse polarization functions of angular momentum $L$ have in analogy with previous work been assigned based on the formula $\zeta_L = (L+1)\zeta_{L=1}$.

For properties depending on the electron density near the nuclei, for example spin−spin coupling constants, the basis-set convergence can be significantly improved by adding tight functions, and we have used the rules established previously[21] for defining pcJ-*n* type basis sets for the He, B, Ne, Al, and Ar elements as well. The augmentation with tight functions for the s-block elements (Li, Be, Na, and Mg) in order to define pcJ-*n* basis sets will have to await a more detailed analysis of the importance of different types of functions for these elements.

## Calibration

The rare gas elements do not form stable molecular species, and it is therefore difficult to evaluate the performance of the

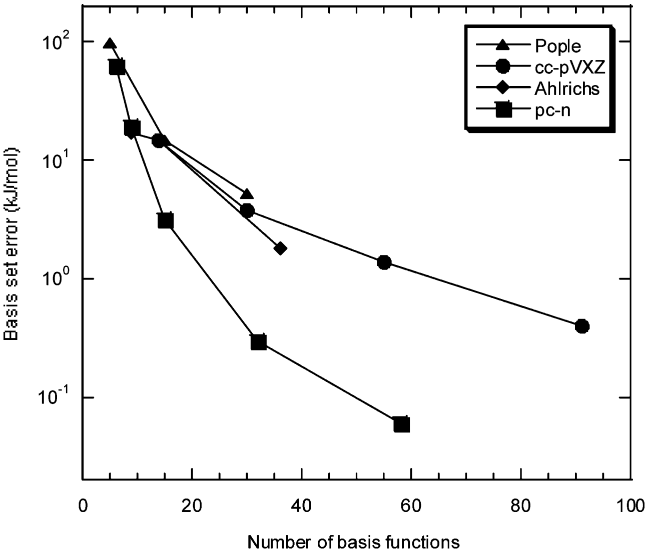

**Figure 3.** Mean absolute deviations (kJ/mol) relative to the basis set limit for the Be systems in Table 1.

pc-*n* basis sets for these elements. The repulsion energy between two rare gas atoms at internuclear distances of 1.4, 1.8, and 2.5 Å (He, Ne, and Ar, respectively) indicated a basis set convergence very similar to the other systems. For the remaining elements, we have selected the systems shown in Table 1 for calibration studies. The performance is evaluated by mean and maximum absolute deviation (MAD and MaxAD) relative to reference values taken as the (uncontracted) pc-2, -3, and -4 extrapolated results.[2]

Table 6 shows the performance for atomization energies for the B and Al systems in Table 1 relative to the diatomic molecular systems, which, as argued earlier,[2] is a better criterion than atomization energies relative to the isolated atoms for density functional methods. The difference between the uncontracted and contracted versions of the pc-*n* basis sets shows that the contraction error in all cases is smaller than the inherent error relative to the basis set limit. The performance of the pc-*n* basis sets is in all cases very similar to those reported earlier for other first and second row elements.[2,3] A comparison with other basis sets of similar quality (Tables 3 and 4), in Table 6, shows that the pc-*n* basis sets in all cases perform better. The pc-1 results, for example, show basis set errors roughly half of that for the popular 6-31G(d,p) basis set, despite the pc-1 having the same number of contracted functions and fewer primitive functions.

Table 7 shows the performance for atomization energies for the Li, Be, Na, and Mg systems in Table 1. The difference between the uncontracted and contracted versions of the pc-*n* basis sets again indicates that the contraction errors are smaller than the inherent errors relative to the basis set limit, with the possible exception of the pc-4 basis set for Na and Mg. The performance of the pc-*n* basis sets is in all cases very similar to those reported earlier and to the B and Al systems in Table 6. The comparison with other basis sets has been grouped slightly different than in Table 6, as the DZP type basis sets 6-31G(d,p) and cc-pVDZ for these elements contain d functions and, therefore, include polarization functions corresponding to a TZP type basis set in the present classification. Similarly, the 6-311G(2df,2pd) and cc-pVTZ contain f functions and, therefore, are of QZP type in the polarization space. Based on the highest angular momentum included in the basis set, Table 7 clearly shows that the pc-*n* basis sets perform better than other alternatives. It is also notable that the pc-1 basis set performs at par or slightly better than the 6-31G(d,p) and cc-pVDZ basis sets, despite the lack of d functions and fewer primitive s and p functions. Comparing 6-31G(d,p) and pc-2 results shows that the latter reduces the basis set errors by a factor of 3, despite the two basis sets having almost the same number of functions. Similarly, the pc-2 results are at par with those from the 6-311G-(2df,2pd) and cc-pVTZ, despite having one less contracted p, d, and f functions, as well as fewer primitive s and p functions (Tables 3 and 4). Figure 3 shows the MAD values for the Be systems as a function of the number of basis functions for the four classes of basis sets. Clearly the pc-*n* basis sets provide a substantial improvement in reducing basis set errors without increasing the size of the basis set.

## Conclusions

The previously proposed methodology for deriving polarization consistent basis sets (pc-*n*) has been applied to the elements He, Li, Be, B, Ne, Na, Mg, Al, and Ar. With these basis sets, there now exist a well-defined hierarchy for approaching the Kohn−Sham basis set limit for molecular calculations for systems composed of the first 18 elements in the periodic table. Basis sets augmented with diffuse functions (aug-pc-*n*) for improving the performance for calculating, e.g., polarization, and augmentation with tight functions (pcJ-*n*) for improving the performance for nuclear spin−spin coupling constants have similarly been defined.

**Acknowledgment.** This work was supported by grants from the Danish Center for Scientific Computation and the Danish Natural Science Research Council.

**Supporting Information Available:** Tables containing exponents and contraction coefficients for the pc-*n* basis sets. This material is available free of charge via the Internet as a link associated with the manuscript.

## References and Notes

(1) Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113; **2002**, *116*, 3502.

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

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

(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, 2002. Jensen, F. *Introduction to Computational Chemistry*; Wiley: New York, 2006.

(5) Dalton, a molecular electronic structure program, experimental version, written by Helgaker, T.; Jensen, H. J. Aa.; 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.; Schimmelpfennig, B.; Sylvester-Hvid, K. O.; Taylor, P. R.; Vahtras, O.

(6) 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 C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(7) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(8) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

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

(10) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *123*, 154101. Becke, A. D.; Johnson, E. R *J. Chem. Phys.* **2006**, *124*, 014104. Johnson, E. R.; Becke, A.D. *J. Chem. Phys.* **2006**, *124*, 174104. Johnson, E. R; Becke, A.D. *Chem. Phys. Lett.* **2006**, *432*, 600.

(11) Petersson, G. A.; Zhong, S.; Montgomery, J. A.; Frisch, M. J. *J. Chem. Phys.* **2003**, *118*, 1101.

(12) The basis sets are provided as supplementary material and will be been made available at the EMSL basis set exchange (http://gnode2.pnl.gov/bse/portal).

(13) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657. Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1970**, *52*, 2769.

(14) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257. Dill, J. D.; Pople, J. A. *J. Chem. Phys.* **1975**, *62*, 2921. Francl, M. M.; Petro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys*. **1982**, *77*, 3654.

(15) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650. McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.

(16) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007. Woon, D. E., Dunning, T. H., Jr. *J. Chem. Phys.* **1994**, *100*, 2975. Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358. The basis sets for Na and Mg are unpublished and have been obtained from the EMSL basis set exchange (http://gnode2.pnl.gov/bse/portal). For the elements Al−Ar, the versions with one additional tight d function have been used, Dunning, T. H., Jr.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys*. **2001**, *114*, 9244.

(17) Schafer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571.

(18) Schafer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829.

(19) Weigend, F.; Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2001**, *119*, 12753.

(20) Jensen, F. *J. Chem. Phys.* **2002**, *117*, 9234. Jensen, F. *J. Chem. Phys.* **2003**, *118*, 2459.

(21) Jensen, F. *J. Chem. Theo. Comp.* **2006**, *2*, 1360.