Theor Chem Acc (2012) 131:1111
DOI 10.1007/s00214-012-1111-4

REGULAR ARTICLE

Electrically polarized valence basis sets for the SBKJC effective core potential developed for calculations of dynamic polarizabilities and Raman intensities

Luciano N. Vidal · Pedro A. M. Vazquez

Received: 2 September 2011 / Accepted: 2 December 2011 / Published online: 4 February 2012

Abstract Sadlej's electric polarization method of Gaussian basis functions was applied to the double-zeta effective core potential basis sets of Stevens, Basch, Krauss, Jasien and Cundari to generate a new augmented polarized valence double-zeta set, named as pSBKJC, which is appropriate for the calculation of dynamic polarizabilities and Raman intensities. The pSBKJC basis set was developed for the atoms of families 14–17 (from C to F, Si to Cl, Ge to Br and Sn to I). In order to assess the performance of this new basis set, these properties were compared to those evaluated using Sadlej's set, available in the EMSL online library under the name of Sadlej-pVTZ. In these tests, Hartree-Fock/pSBKJC calculations have proved to be less demanding of the computer than the Hartree-Fock/Sadlej-pVTZ ones but give results in excellent agreement with those from the Sadlej-pVTZ basis set. Since the Stevens et al. pseudopotential can represent the scalar relativistic effects, the results obtained at the Hartree-Fock/pSBKJC level show a better agreement with the results of Dirac-Hartree-Fock/Sadlej-pVTZ relativistic calculations using Dyall's spin-free Hamiltonian. When comparing Hartree-Fock/pSBKJC data of Raman scattering activities, at the excitation wavelength of 488 nm, with those of spin-free Dirac-Hartree-Fock/Sadlej-pVTZ calculations, a very good agreement is observed, where the RMS error is 8.5 $\mathring{A}^4$a.m.u.$^{-1}$ and the averaged percentage error is 3.4%. In terms of computer savings in calculations of dynamic Raman intensities, a 20% reduction in the CPU time in the coupled cluster singles and doubles intensities of C_6H_6 and about 40% reduction in the time-dependent Hartree-Fock intensities for C_6F_6 molecules were attained.

Keywords Raman spectroscopy · Static and dynamical polarizabilities · Relativistic effects · Ab initio electronic structure · ECP basis set

1 Introduction

In the Raman effect, the scattering intensity depends on the frequency ν_{ex} of the excitation light. The intensity of each Raman transition usually varies if ν_{ex} is changed, and the relative values of the Raman intensities can also depend on ν_{ex} [1, 2]. This behavior is explained through the theory of light scattering by the ν_{ex}^4 factor, since the scattered electromagnetic radiation originates from an oscillating electric dipole, appearing in the expression of the cross section and by the presence of ν_{ex} in the polarizability formula itself [3, 4]. In the static limit approximation, where $\nu_{ex} = 0$, the Raman intensities are evaluated with a computational cost similar to that of the calculation of harmonic frequencies [5] and many electronic structure codes usually adopt this approximation to compute the Raman intensities [6, 7]. However, within the static limit ,many aspects of the Raman spectrum are not observed, like the *excitation profile*—the dependence of the intensity of each Raman transition on ν_{ex}. In recent years, electronic structure methods were reported and implemented for the calculation of the *dynamic polarizabilities* at the HF, MCSCF, MP2, Coupled cluster and DFT levels [8–12], enabling the calculation of Raman intensities without the restriction of

L. N. Vidal (✉)
Chemistry and Biology Department, Federal Technological University of Paraná, Curitiba, PR, Brazil
e-mail: lnvidal@utfpr.edu.br

P. A. M. Vazquez
Physical-Chemistry Department, Chemistry Institute, State University of Campinas, Campinas, SP, Brazil
e-mail: vazquez@iqm.unicamp.br

$\nu_{ex} = 0$ by finite difference methods [13–17] or even from analytic derivatives [18, 19].

In studies concerning the absolute values of Raman cross sections, the electron correlation effects [13–17, 20], basis set convergence [17], relativistic [21, 22] and solvation effects [23] were investigated. These papers have shown the need of frequency-dependent polarizability gradients at the CCSD/aug-cc-pVTZ level for quantitative agreement between theoretical and gas phase experimental data. However, such level of theory is very demanding of computer time, thus being limited to the study of small-sized systems. As an alternative to reduce the computational needs in these calculations, the spatial symmetry can be explored in the numerical differentiation of the polarizabilities (the most expensive step) [15, 17]. Nevertheless, since the majority of molecules are non-symmetrical, other alternatives must be considered.

In electronic structure calculations, the computational cost is strongly dependent on the number of basis functions and significant savings are obtained with the use of purpose-oriented basis sets. Following this strategy, Sadlej developed medium-size polarized basis sets [24–27], which give the best relation between the quality of the Raman intensities and the computational cost. Using Sadlej's basis, named as Sadlej-pVTZ in the EMSL basis set library [28, 29], the dynamic polarizabilities and Raman cross sections are evaluated with similar quality to the aug-cc-pVTZ set [17, 30] but with an order of magnitude reduction in the computer time [17, 31].

The importance of inner electrons on the dynamic Raman intensities was assessed in the CCSD [14, 16] and CC3 [20] levels with the aug-cc-pCVTZ basis where it was shown that the excitation of these electrons has little influence on the Raman properties (cross section and scattering activity). Considering such evidence, this paper proposes that the inner electrons can be removed from the atoms and represented by an effective core potential (ECP) without loss of quality of the Raman properties.

Our approach for reduction in computer requirements in polarizability and Raman intensity calculations is the use of the *basis set polarization method*, developed by Sadlej [24] and applied in the generation of the Sadlej-pVTZ set, to develop a new basis set for calculations of dynamic polarizabilities and Raman intensities. This new set, the *pSBKJC*,[1] derives from the relativistic pseudopotential basis set of Stevens and collaborators [32–34] where the ECP were modeled to take into account the scalar relativistic effects (mass-velocity and Darwin corrections) from numerical atomic Dirac-Fock calculations. Therefore, as will be shown, the pSBKJC set is able to represent both the electric polarization and the scalar relativistic effects on the polarizabilities and Raman intensities but is less computer demanding than the Sadlej-pVTZ set.

In the following section, the *Gaussian basis set electric polarization method* is briefly described and the expressions for the properties treated herein are presented. The next section provides the computational details, and then, the results obtained with this new basis set are presented and discussed.

2 Theory

The polarization procedure first requires the selection of a basis set that must adequately describe the unperturbed system. For this purpose, the valence double-zeta pseudopotential basis set of Stevens and coworkers was chosen [32–34].

The polarized basis sets are obtained from analysis of the dependency of a spin-orbital u_i of a many-electron system on the perturbation parameter λ associated with a homogeneous electric field $\mathbf{F}$ interacting with the electric dipole moment $\mathbf{p}$ of the system. First-order perturbation theory is used to describe this interaction, whose total Hamiltonian H is

$$H = H^{(0)} - \lambda \mathbf{F} \cdot \mathbf{p} \tag{1}$$

in which $H^{(0)}$ and $-\lambda \mathbf{F} \cdot \mathbf{p}$ are the zeroth and the first-order components of H from the perturbed system. The perturbed spin-orbitals $u_i(\lambda)$ are written as a linear combination of basis functions χ_μ, explicitly dependent on the external perturbation,

$$u_i(\lambda) = \sum_\mu c_{i\mu}(\lambda)\chi_\mu(\lambda) \tag{2}$$

The rule to determine $u_i(\lambda)$ is obtained by applying some restrictions to the perturbed system. The wave-function of the system is represented by a single Slater determinant whose spin-orbitals are found by solving the Hartree-Fock equation, approaching $c_{i\mu}(\lambda)$ to $c_{i\mu}(0)$ and assuming that χ_μ are given by Gaussian-type orbitals (GTO). The final form of $u_i(\mathbf{F})$, the polarized spin-orbital, is

$$u_i(\mathbf{F}) = \sum_\mu c_{i\mu}(0)\chi_\mu^l(\mathbf{F} = \mathbf{0}) + \sum_\mu c_{i\mu}(0)\alpha_\mu^{-1/2}\chi_\mu^{l+1}(\mathbf{F} = \mathbf{0}) \tag{3}$$

where $\chi_\mu^l(\mathbf{F} = \mathbf{0})$ is a GTO of the unperturbed system ($\mathbf{F} = \mathbf{0}$) with its respective linear combination coefficient $c_{i\mu}(0)$. The superscript "l" refers to the angular momentum part of the GTO. The electric polarization of the spin-orbital u_i is introduced by the second sum in Eq. 3 where $\chi_\mu^{l+1}(\mathbf{F} = \mathbf{0})$ is a new GTO with same exponent α_μ of

[1] Following the nomenclature from the EMSL basis set library, the Stevens and coworkers bases are called here as the SBKJC set and the prefix "p" stands for *polarized*.

$\chi_{\mu}^{l}(\mathbf{F}=\mathbf{0})$ but the angular function is increased by one unit in the angular momentum quantum number l.

After some exploratory calculations of the polarizability for some atomic systems, Sadlej proposed a set of five steps to generate the polarized basis set. These steps are summarized below:

1. The original double-zeta basis set is augmented by one primitive diffuse function for each shell. Their exponents are derived from the corresponding even-tempered sequences.
2. The polarization functions are generated only for the outermost occupied shell of the given atom. The contraction coefficient in Eq. 3 is determined from Hartree-Fock eigenvectors of the nearest negative ion.
3. Only the four most diffuse polarization functions are retained in the basis. These are contracted on the form: $4d \rightarrow 2d + 2d$.

The method of electric polarization of Gaussian basis functions was applied to the valence double-zeta pseudo-potential basis sets SBKJC to generate a new basis set, named as *pSBKJC*, which is suitable for calculations of dynamic polarizabilities and Raman cross sections, as will be shown in Sect. 4. The pSBKJCs are augmented polarized double-zeta basis sets. The number of primitives and contracted basis functions of the sets pSBKJC and Sadlej-pVTZ (given for comparison) are shown in Table 1.

2.1 Raman cross section and scattering activity

The differential Raman Stokes cross section corresponding to the fundamental intensity of the k-th vibrational normal mode, measured perpendicular to the incident light at some temperature T and collected within the solid angle Ω, is given by [3]:

$$\left(\frac{d\sigma}{d\Omega}\right)_k = \frac{(2\pi)^4}{45}\frac{(\tilde{\nu}_{ex}-\tilde{\nu}_k)^4}{(1-\exp(-hc\tilde{\nu}_k/k_BT))}\frac{h}{8\pi^2 c\tilde{\nu}_k}S_k \qquad (4)$$

with $\tilde{\nu}_{ex}$ and $\tilde{\nu}_k$ being the excitation light and the normal vibration wavenumbers and S_k being the Raman Scattering Activity. The above expression is valid for linearly polarized incident light with polarization perpendicular to the scattering plane and scattered light observed without the use of polarizers. The constants h, c and k_B have their usual meaning. The Raman scattering activity is defined as follows:

$$S_k \equiv \left[45\left(\frac{\partial\bar{\alpha}}{\partial Q_k}\right)^2_{eq} + 7\left(\frac{\partial\gamma}{\partial Q_k}\right)^2_{eq}\right]g_k \qquad (5)$$

where $\bar{\alpha}$ and γ are the mean polarizability and the polarizability anisotropy, respectively, Q_k is the kth normal coordinate and g_k the degenerescence of the mode k. Both $\bar{\alpha}$ and γ depend on the frequency of the excitation light, *that is*, they are dynamic properties. The Eqs. 4 and 5 are used to express intensity values of Raman scattering and both are used in this paper.

Table 1 Size and composition of the basis sets pSBKJC and Sadlej-pVTZ

Atoms	pSBKJC		Sadlej-pVTZ	
	Primitives	Contracted	Primitives	Contracted
C–F	5s.5p.4d	3s.3p.2d	10s.6p.4d	5s.3p.2d
Si–Cl	5s.5p.4d	3s.3p.2d	13s.10p.4d	7s.5d.2d
Ge–Br	5s.5p.4d	3s.3p.2d	15s.12p.9d	9s.7p.4d
Sn–I	5s.5p.4d	3s.3p.2d	19s.15p.12d	11s.9p.6d

3 Computational details

The Hartree-Fock/Sadlej-pVTZ calculations were performed with the electronic structure program DALTON [35] and the Dirac-Hartree-Fock (Spin-Free Hamiltonian)/Sadlej-pVTZ and Dirac-Hartree-Fock (Dirac-Coulomb Hamiltonian)/Sadlej-pVTZ calculations by the relativistic ab initio code DIRAC [36]. The acronyms HF for Hartree-Fock, DHF-SF for Dirac-Hartree-Fock (Spin-Free Hamiltonian) and DHF-DC for Dirac-Hartree-Fock (Dirac-Coulomb Hamiltonian) are used from now on when referring to these methods/Hamiltonians. Geometry optimizations and quadratic force constants were evaluated at the HF/Sadlej-pVTZ level. Dynamic polarizabilities were calculated using the linear response (LR) modules of the DALTON and DIRAC programs but the HF/pSBKJC frequency-dependent polarizabilities were obtained from time-dependent HF calculations using the GAMESS [6] electronic structure program. The Raman cross sections and scattering activities were evaluated by finite difference procedures using our Fortran 77 code PLACZEK [16, 17].

In order to eliminate the influence of the molecular geometry in the results presented here, we decided to use the same geometry for both basis sets (Sadlej-pVTZ and pSBKJC) in all polarizabilities and Raman intensities calculations. These geometries were evaluated at the HF/Sadlej-pVTZ level. For a similar reason, the normal coordinates and harmonic frequencies were kept fixed at the HF/Sadlej-pVTZ level. In the calculations with the pSBKJC set, the Sadlej-pVTZ basis was used in the hydrogen atoms.

4 Results and discussion

The pSBKJC bases were developed for the atoms of groups 14 (C–Sn), 15 (N–Sb), 16 (O–Te) and 17 (F–I) where the double-zeta valence basis set SBKJC [32–34] was electrically polarized following the procedure described in the Theory section. The pseudopotential SBKJC for these atoms keeps the valence *s* and *p* subshell electrons (e.g., $5s^2 5p^5$ for iodine) and replaces the inner ones by an energy-consistent pseudopotential modeled to represent the scalar relativistic effects (mass-velocity and Darwin). The quality of the polarizabilities and the Raman intensities computed using the pSBKJC set was assessed by taking as reference data the values obtained using the Sadlej-pVTZ set. These properties were calculated at the HF, DHF-SF and DHF-DC levels with Sadlej-pVTZ and pSBKJC basis sets for the followings molecules: XH_4 (X = C, Si, Ge or Sn), XH_3 (X = N, P, As or Sb), H_2X (X = O, S, Se or Te) and HX (X = F, Cl, Br or I). To assess the results from the new basis set, the function $\delta_{\%}$, which gives the percentage deviation between two sets of data weighted by the values of the reference data, was defined. If the property being assessed is S_k then $\delta_{\%}$ is given by:

$$\delta_{\%}(S_k) = \frac{\sum_k |S_k(\lambda_{ex}, \text{Sadlej}-\text{pVTZ}) - S_k(\lambda_{ex}, \text{pSBKJC})|}{\sum_k S_k(\lambda_{ex}, \text{Sadlej-pVTZ})} \times 100\% \tag{6}$$

in which $S_k(\lambda_{ex}$, Sadlej-pVTZ) is the Raman scattering activity of the *k*th normal vibration evaluated using the Sadlej-pVTZ set at the excitation wavelength λ_{ex} and $S_k(\lambda_{ex}$, pSBKJC) has an analogous meaning. The function $\delta_{\%}$ is calculated by summing over all Raman active modes of all molecules.

The percentage deviation $\delta_{\%}$ computed for $\bar{\alpha}$ and γ for the sixteen molecules listed above at several excitation wavelengths is presented in Table 2. For the mean polarizabilities, $\delta_{\%}$ is very small varying from 1.2 to 1.4%, and the agreement between the pSBKJC data and the relativistic data sets is slightly better. On the other hand, the agreement observed for γ is very poor, where $\delta_{\%}$ is about 30%. By comparing the values of $\bar{\alpha}$ and γ given in Tables 3 and 4, we see that the absolute errors of the pSBKJC's are of the order of 10^{-1}bohr^3, with respect to the Sadlej-pVTZ data. Within this precision, the new basis set is not

Table 2 Percentage deviation (function $\delta_{\%}$, Eq. 6) for the polarizabilities $\bar{\alpha}$ and γ at several excitation wavelengths λ_{ex}

	λ_{ex}/nm			
	Static	632.8	514.5	488.0
$\delta_{\%}$ of $\bar{\alpha}$				
HF/Sadlej-pVTZ	1.4	1.4	1.4	1.4
DHF-SF/Sadlej-pVTZ	1.2	1.2	1.2	1.2
DHF-DC/Sadlej-pVTZ	1.2	1.3	1.3	1.3
$\delta_{\%}$ of $\vert\gamma\vert$				
HF/Sadlej-pVTZ	31.8	30.5	29.2	28.4
DHF-SF/Sadlej-pVTZ	32.6	32.6	32.1	31.7
DHF-DC/Sadlej-pVTZ	31.5	31.3	30.7	30.7

Comparison between HF/pSBKJC results with three different Sadlej-pVTZ sets of data

Table 3 Mean polarizabilities $\bar{\alpha}$ (in bohr^3) for λ_{ex} = 488.0 nm

	HF	HCl	HBr	HI
HF/Sadlej-pVTZ	4.878	17.262	24.043	37.367
DHF-SF/Sadlej-pVTZ	4.885	17.291	24.073	36.991
DHF-DC/Sadlej-pVTZ	4.885	17.292	24.104	37.189
HF/pSBKJC	4.877	17.154	23.815	35.945
	H_2O	H_2S	H_2Se	H_2Te
HF/Sadlej-pVTZ	8.555	25.072	31.979	46.595
DHF-SF/Sadlej-pVTZ	8.565	25.111	31.963	46.211
DHF-DC/Sadlej-pVTZ	8.565	25.112	31.984	46.339
HF/pSBKJC	8.611	25.019	31.829	47.094
	NH_3	PH_3	AsH_3	SbH_3
HF/Sadlej-pVTZ	13.235	31.786	36.973	50.887
DHF-SF/Sadlej-pVTZ	13.245	31.803	36.906	50.122
DHF-DC/Sadlej-pVTZ	13.245	31.804	36.913	50.167
HF/pSBKJC	13.025	31.447	36.464	49.111
	CH_4	SiH_4	GeH_4	SnH_4
HF/Sadlej-pVTZ	16.482	31.547	34.916	44.539
DHF-SF/Sadlej-pVTZ	16.488	31.595	35.277	45.607
DHF-DC/Sadlej-pVTZ	16.488	31.595	35.282	45.640
HF/pSBKJC	16.599	31.338	35.297	45.090

Table 4 Absolute values of the polarizability anisotropy $|\gamma|$ (in bohr^3) for λ_{ex} = 488.0 nm

	HF	HCl	HBr	HI
HF/Sadlej-pVTZ	1.188	1.947	1.914	2.244
DHF-SF/Sadlej-pVTZ	1.191	1.939	1.908	2.270
DHF-DC/Sadlej-pVTZ	1.191	1.940	1.925	2.340
HF/pSBKJC	1.204	2.012	2.549	3.523
	H_2O	H_2S	H_2Se	H_2Te
HF/Sadlej-pVTZ	0.971	0.709	1.740	2.790
DHF-SF/Sadlej-pVTZ	0.973	0.732	1.775	2.814
DHF-DC/Sadlej-pVTZ	0.973	0.731	1.759	2.702
HF/pSBKJC	0.952	0.654	0.933	2.744
	NH_3	PH_3	AsH_3	SbH_3
HF/Sadlej-pVTZ	0.973	1.370	1.385	0.403
DHF-SF/Sadlej-pVTZ	0.976	1.403	1.437	0.038
DHF-DC/Sadlej-pVTZ	0.976	1.402	1.429	0.076
HF/pSBKJC	0.730	0.966	0.611	1.075

Table 5 Percentage deviations (function $\delta_{\%}$, Eq. 6) for Raman scattering activities S_k at several excitation wavelengths λ_{ex}

	λ_{ex} / nm			
	Static	632.8	514.5	488.0
$\delta_{\%}$ of S_k				
HF/Sadlej-pVTZ	6.6	6.7	6.7	6.8
DHF-SF/Sadlej-pVTZ	3.8	3.6	3.5	3.4
DHF-DC/Sadlej-pVTZ	3.9	3.7	3.6	3.6

Comparison between HF/pSBKJC results with three different Sadlej-pVTZ sets of data

appropriated for calculations of polarizabilities which are less than unit, like γ of these test molecules. The percentage deviation evaluated for the Raman scattering activities is collected in Table 5. The data in this table show that $\delta_{\%}$ is almost independent of λ_{ex}. Thus, the pSBKJC set performs as good as the Sadlej-pVTZ one in calculations of static and dynamic Raman intensities (here expressed by S_k) with an agreement falling within the interval of 3.4–6.8%, depending on the method or the λ_{ex} considered for the reference data. An expected and confirmed result is that the agreement becomes better if we compare pSBKJC results with relativistic data, especially with spin-free (DHF-SF/Sadlej-pVTZ) values of S_k. Since the SBKJC pseudopotentials were modeled to also represent the scalar relativistic effects, the agreement is about twice as good when pSBKJC data are compared with spin-free values of S_k. From the computational point of view, these results are important as the scalar (SF Hamiltonian) or full (DC) relativistic calculations are much more expensive than their non-relativistic counterpart (HF level). The treatment of the relativistic effects can also be done by less computer demanding methods, like the Douglas-Kroll-Hess Hamiltonian [37–39], which have a computational cost equivalent to the non-relativistic case but the advantage of the use of ECP basis sets is still evident due to the small number of basis functions for the heavier atoms. To illustrate the case to case accordance of the scattering activities, the HF/pSBKJC S_k at $\lambda_{ex} = 488$ nm of all molecules were plotted in Fig. 1 against the three sets of Sadlej-pVTZ data. This figure shows that the agreement is very satisfactory between these data, especially with those from relativistic calculations. The RMS error computed for these three pairs of data are 21.1, 8.5 and 9.3 Å^{4}a.m.u.$^{-1}$, respectively, for HF, DHF-SF and DHF-DC/Sadlej-pVTZ. Again, the overall accordance is better with scalar relativistic values of Raman intensities.

The pSBKJC basis set was also used in correlated calculations where the CCSD differential Raman scattering cross sections of C_6H_6 and C_6D_6 were calculated at $\lambda_{ex} = 488$ nm, using the pSBKJC and Sadlej-pVTZ sets. The inner electrons were kept frozen in the calculations

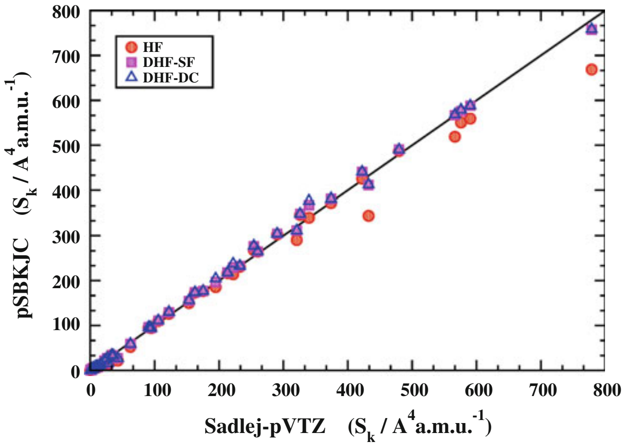

Fig. 1 Comparison of the Raman scattering activities (S_k in Å^{4}a.m.u.$^{-1}$) calculated with the pSBKJC basis set against the data from the Sadlej-pVTZ set. Excitation wavelength is 488 nm

with the Sadlej-pVTZ set. The computed values of these cross sections are shown in Table 6. In these systems, the pseudopotential removes 12 electrons, and the total number of contracted basis functions is 198 for Sadlej-pVTZ and 186 for pSBKJC. The gradients of $\bar{\alpha}$ and γ are evaluated numerically in PLACZEK, which is able to explore the molecular symmetry in a very efficient way. Thus, instead of 73 single-point calculations, only 11 were necessary to take the fundamental Raman cross sections of these molecules. When pSBKJC was used for the carbon atoms, this time was reduced by 20%. As can be seen in Table 6, this

Table 6 Differential Raman scattering cross sections of C_6H_6 and C_6D_6 (in 10^{-36}m^{2}sr^{-1}) for the excitation wavelength of 488.0 nm at 300 K

Mode	Type	$\tilde{\nu}$/cm^{-1} [45]	Experimental [46]	CCSD/	
				Sadlej-pVTZ	pSBKJC
C_6H_6					
$\nu_1(a_{1g})$	ν_{CH}	3,062	644 ± 167	675.1	674.6
$\nu_2(a_{1g})$	ν_{ring}	992	710 ± 71	902.5	880.7
$\nu_{11}(e_{1g})$	ω_{CH}	849	25 ± 5	37.3	47.8
$\nu_{15}(e_{2g})$	ν_{CH}	3,047	439 ± 194	420.3	427.1
$\nu_{16}(e_{2g})$	ν_{ring}	1,596	129 ± 26	181.8	187.6
$\nu_{17}(e_{2g})$	δ_{CH}	1,178	66 ± 13	54.8	55.2
$\nu_{18}(e_{2g})$	δ_{ring}	606	76 ± 15	115.6	108.8
C_6D_6					
$\nu_1(a_{1g})$	ν_{CD}	2,293	372 ± 92	427.8	428.7
$\nu_2(a_{1g})$	ν_{ring}	943	607 ± 61	973.3	951.5
$\nu_{11}(e_{1g})$	ω_{CD}	662	53 ± 11	84.7	105.7
$\nu_{15}(e_{2g})$	ν_{CD}	2,265	244 ± 113	284.4	289.6
$\nu_{16}(e_{2g})$	ν_{ring}	1,552	114 ± 23	186.4	192.5
$\nu_{17}(e_{2g})$	δ_{CD}	867	91 ± 18	97.0	97.2
$\nu_{18}(e_{2g})$	δ_{ring}	577	65 ± 13	105.2	98.8

significant reduction in computer time is accompanied by not much deterioration of the Raman cross sections. For all but the $\nu_{11}(e_{1g})$ mode, the percentage deviation with respect to the Sadlej-pVTZ data varies from zero to 6% and for ν_{11} is 28% and 25% for benzene and deutered benzene, respectively. The average deviation between CCSD/Sadlej-pVTZ and experimental data is 34% and for CCSD/pSBKJC and experimental is 38%. These results agree only semi-quantitatively with the experimental data available as the average experimental uncertainty for these molecules is 23%. In a previous study about Raman intensities, it was shown that the CCSD/Sadlej-pVTZ and CCSD/aug-cc-pVQZ data differ about 5.6% (S_{abs} in Table IV of ref. [17]). This suggests that these large deviations are not only related to the lack of flexibility of the pSBKJC and Sadlej's bases but can also originate from the need of more sophisticated treatments of the electron correlation and from anharmonic effects.

Coupled cluster calculations were also done for the following molecules: CH_3F, NH_3F, H_2O, SiH_3Cl, PCl_3, H_2S, GeH_3Br, $AsBr_3$, H_2Se, SnH_4, SbH_3 and H_2Te. All normal modes of these twelve molecules are Raman active giving a total of fifty fundamental transitions. The reference data are frozen-core CCSD/Sadlej-pVTZ calculations at $\lambda_{ex} =$ 488 nm. Following a procedure similar to that described in section 3, the Sadlej-pVTZ basis was used in the hydrogen atoms and the CCSD/Sadlej-pVTZ and CCSD/pSBKJC Raman intensities were evaluated using geometries and harmonic force fields at the CCSD/Sadlej-pVTZ level. The comparison of these data is shown in Fig. 2 where the CCSD/pSBKJC scattering activities are plotted against the CCSD/Sadlej-pVTZ results. The agreement between these data is in general very good with the largest difference occurring for S_k of the mode $\nu_1(a_1)$ of SnH_4, for which the CCSD/Sadlej-pVTZ result is 714 Å^4a.m.u.$^{-1}$ and the CCSD/pSBKJC is 829 Å^4a.m.u.$^{-1}$ (16% of difference). The values of the function $\delta_{\%}$ computed for $\bar{\alpha}$, γ and S_k are 3.1, 8.3 and 7.9%, respectively. The CPU time required to calculate the Raman intensities for all these molecules using the pSBKJC basis sets is 82% of that needed to perform the Sadlej-pVTZ calculations. Therefore, on average, the use of the pSBKJC resulted in Raman intensities 7.9% different from the reference data but demanding 18% less computer time, showing the advantage of using the pSBKJC basis set in correlated calculations.

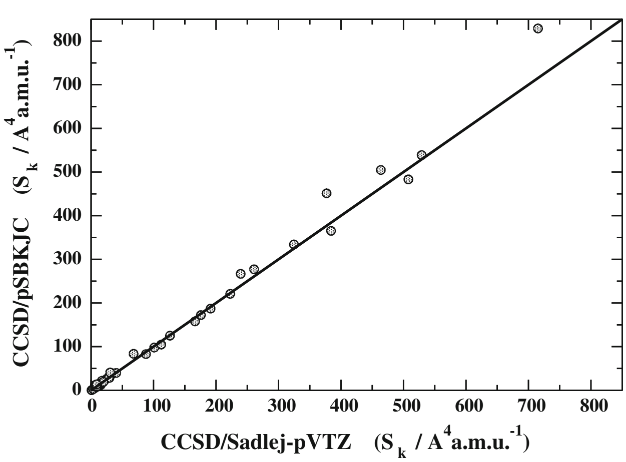

Fig. 2 Comparison of the CCSD/pSBKJC Raman scattering activities (S_k in Å^4a.m.u.$^{-1}$) with the CCSD/frozen-core/Sadlej-pVTZ data. Excitation wavelength is 488 nm

The polarizabilities and the differential Raman cross sections for the C_6F_6 molecule are presented in Tables 7 and 8. In this case, where 24 electrons are dropped by the pseudopotential, the computational cost measured in terms of CPU time (see Table 7) is about 40% less when pSBKJC is used. The number of contracted basis functions for C_6F_6 is 288 and 264 for Sadlej-pVTZ and pSBKJC, respectively. As the linear response calculations are very dependent on the number of electrons and basis functions, an apparently small reduction in these number (24 less electrons and contracted basis sets) results in a great reduction in CPU time. Due to the numerical approach used the get the polarizability derivatives, the time savings for the Raman intensity calculations is very close to this value (since non-equilibrium geometries are only 10^{-4} bohr away from equilibrium). The values for the polarizabilities and Raman cross sections presented in Tables 7 and 8 show a very good agreement with the Sadlej-pVTZ results. In Table 8, the cross sections obtained from a calculation using SBKJC set were also included. By comparing the cross sections themselves or the RMS errors given in this table, we see that the electric polarization of the SBKJC basis substantially improves the values of the Raman intensities. The RMS errors of SBKJC and pSBKJC, with respect to Sadlej-pVTZ data, are 31.6 and 5.8 Å^4a.m.u.$^{-1}$, respectively. In Table 8, the Raman cross sections evaluated using the Z2Pol basis set are also given. The Z2Pol set [40] was developed by Sadlej and coworkers using the electric polarization method, aiming at a reduction in the computer requirements of electric properties calculations. For C_6F_6, the number of contracted Z2Pol basis functions is 216 but all electrons are kept (no pseudopotential). As the number of basis functions is significantly lower than those of Sadlej-pVTZ or even pSBKJC, the total CPU time for a

Table 7 Computer CPU time and polarizabilities (in bohr^3) for the molecule C_6F_6

	Relative CPU time	λ_{ex} static		$\lambda_{ex} = 488.0$ nm	
		$\bar{\alpha}$	$\lvert\gamma\rvert$	$\bar{\alpha}$	$\lvert\gamma\rvert$
HF/pSBKJC	0.6	64.609	36.725	67.241	39.351
HF/Sadlej-pVTZ	1.0	65.101	36.989	67.758	39.641

Table 8 Differential Raman scattering cross sections of C_6F_6 (in $10^{-36}m^2sr^{-1}$) for the excitation wavelength of 488.0 nm and temperature of 300 K

Mode	$\tilde{\nu}/cm^{-1}$ [47]	HF/			
		Sadlej-pVTZ	pSBKJC	SBKJC	Z2Pol
$\nu_1(a_{1g})$	1,490	54.4	57.2	10.3	37.7
$\nu_2(a_{1g})$	559	535.2	521.0	554.7	565.6
$\nu_{11}(e_{1g})$	370	288.7	289.3	328.1	265.3
$\nu_{15}(e_{2g})$	1,655	44.7	42.9	69.6	50.8
$\nu_{16}(e_{2g})$	1,157	17.9	19.0	43.0	23.6
$\nu_{17}(e_{2g})$	443	98.7	94.1	141.3	102.3
$\nu_{18}(e_{2g})$	264	4.4	4.9	9.7	4.6
RMS[a]			5.8	31.6	16.2

[a] Taking as reference the Sadlej-pVTZ set

linear response HF calculation with the Z2Pol set is 47% less than the time of the HF/Sadlej-pVTZ level. From the data in Table 8, the quality of the differential Raman cross sections of Z2Pol is intermediate between Sadlej-pVTZ and SBKJC intensities, where the RMS error for the Z2Pol is 16.2 $Å^4a.m.u.^{-1}$. Thus, Z2Pol represents an alternative for computation of Raman intensities of larger systems with reasonable accuracy.

As a final concern, we wish to point out that standard and augmented correlation consistent-like basis sets, employing relativistic pseudopotentials, from double to quintuple-zeta size have been reported for group 13-18 elements (Ga-Rn) and also for 4d and 5d elements [41–44]. The so-called cc-pVnZ-PP and aug-cc-pVnZ-PP pseudopotential correlation consistent basis sets provided good results for equilibrium bond lengths, vibrational frequencies and dissociation energies of diatomic systems in CCSD(T) calculations but have not yet been employed in polarizability and Raman intensity calculations. Since the aug-cc-pVTZ performs very well for polarizability and Raman intensities, it would be valuable to assess the aug-cc-pVTZ-PP set in such property calculations.

5 Summary and conclusions

Starting from the valence double-zeta pseudopotential basis SBKJC from Stevens and coworkers [32–34], Sadlej's electric polarization method [24] was used to generate a new set called pSBKJC, which is suitable for calculations of static and dynamic polarizabilities and Raman intensities. The results obtained at the HF/pSBKJC level for the molecules XH_4 (X = C, Si, Ge or Sn), XH_3 (X = N, P, As or Sb), H_2X (X = O, S, Se or Te) and HX (X = F, Cl, Br or I) were compared with three sets of data: non-relativistic HF/Sadlej-pVTZ and relativistic DHF(spin-free Hamiltonian)/Sadlej-pVTZ and DHF(Dirac-Coulomb Hamiltonian)/Sadlej-pVTZ. Since the SBKJC pseudopotential can represent the scalar relativistic effects, a better agreement was observed with the results of spin-free DHF/Sadlej-pVTZ calculations, where the RMS error in the Raman scattering activities is 8.5 $Å^4a.m.u.^{-1}$ (less than 4%). A 20% reduction in the CPU time in the CCSD calculation of the dynamic Raman cross sections of C_6H_6 and about 40% reduction in the HF calculation for the C_6F_6 molecule were attained. When comparing the CPU time between HF/pSBKJC and DHF/Sadlej-pVTZ calculations, the savings are much larger since the relativistic four-component electronic structure calculations are much more demanding in computer time then the HF level.

Finally, this study leads also to the conclusion that for the systems studied here the direct contribution of the inner electrons to the polarizability and Raman intensity is very small, thus enabling the use of pseudopotential approaches to treat both properties. Due to the large computer requirements needed to perform dynamic Raman intensity calculations, the pSBKJC basis set provides an alternative that enables the study of systems containing a large number of electrons and/or heavy nucleus.

Acknowledgments The authors thank the National Center for High Performance Computing in São Paulo (CENAPAD-SP) for computer time. LNV thanks the National Council for Scientific and Technological Development (CNPq) for a doctoral fellowship. The basis sets can be obtained directly from the authors: contact LNV at lnvidal@utfpr.edu.br or PAMV at vazquez@iqm.unicamp.br.

References

1. Asher SA, Johnson CR (1985) J Phys Chem 89:1375
2. de Miranda SG, Vazquez PAM (2002) J Braz Chem Soc 13:324
3. Long DA (2002) The Raman effect: a unified treatment of the theory of Raman scattering by molecules. Wiley, LTD, Chichester
4. Placzek G (1959) The Rayleigh and Raman scattering. United States Atomic Energy Commision, Lawrence Radiation Laboratory, University of California, Livermore, California, UCRL Translation No. 526 (L), Physics
5. Komornicki A, McIver Jr JW (1979) J Chem Phys 70:2014
6. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347
7. Gaussian 98, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Jr., Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Salvador P, Dannenberg JJ, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Baboul AG, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (1998) Gaussian, Inc., Pittsburgh PA, 2001

8. Olsen J, Jørgensen P (1985) J Chem Phys 82:3235
9. Jørgensen P, Jensen HJA, Olsen J (1988) J Chem Phys 89:3654
10. Christiansen O, Halkier A, Koch H, Jørgensen P, Helgaker T (1998) J Chem Phys 108:2801
11. Hättig C, Christiansen O, Jørgensen P (1997) J Chem Phys 107:10592
12. Sałek P, Vahtras O, Helgaker T, Ågren H (2002) J Chem Phys 117:9630
13. Caillie CV, Amos RD (2000) Phys Chem Chem Phys 2:2123
14. Pecul M, Rizzo A (2002) J Chem Phys 116:1259
15. Neugebauer J, Reiher M, Hess BA (2002) J Chem Phys 117:8623
16. Vidal LN, Vazquez PAM (2003) Quim Nova 26:507
17. Vidal LN, Vazquez PAM (2005) Int J Quantum Chem 103:632
18. Quinet O, Champagne B (2001) J Chem Phys 115:6293
19. O'Neill DP, Kállay M, Gauss J (2007) Mol Phys 105:2447
20. Pecul M, Coriani S (2002) Chem Phys Lett 355:327
21. Pecul M, Rizzo A (2003) Chem Phys Lett 370:578
22. Vidal LN, Vazquez PAM (2006) Chem Phys 321:209
23. Corni S, Cappelli C, Cammi R, Tomasi J (2001) J Phys Chem A 105:8310
24. Sadlej AJ (1988) Collect Czech Chem Commun 53:1995
25. Sadlej AJ (1991) Theor Chim Acta 79:123
26. Sadlej AJ (1992) Theor Chim Acta 81:45
27. Sadlej AJ (1992) Theor Chim Acta 81:339
28. Feller D (1996) J Comput Chem 17:1571
29. Schuchardt KL, Didier BT, Elsethagen T, Sun LS, Gurumoorthi V, Chase J, Li J, Windus TL (2007) J Chem Inf Model 47:1045
30. Vidal LN (2004) Master Thesis. State University of Campinas, Campinas, SP, Brazil
31. Oakes RE, Bell SEJ, Benkova Z, Sadlej AJ (2005) J Comput Chem 26:154
32. Stevens WJ, Basch H, Krauss M (1984) J Chem Phys 81:6026
33. Stevens WJ, Krauss M, Basch H, Jasien PG (1992) Can J Chem 70:612
34. Cundari TR, Stevens WJ (1993) J Chem Phys 98:5555
35. (2005) Dalton, a molecular electronic structure program. Release 2.0, see http://www.kjemi.uio.no/software/dalton/dalton.html
36. Jensen HJA, Saue T, Visscher L, with contributions from Bakken V, Eliav E, Enevoldsen T, Fleig T, Fossgaard O, Helgaker TU, Lærdahl J, Larsen CV, Norman P, Olsen J, Pernpointner M, Pedersen JK, Ruud K, Sałek P, van Stralen JNP, Thyssen J, Visser O, Winther T (2004) DIRAC, a relativistic ab initio electronic structure program. Release DIRAC04.0, see http://dirac.chem.sdu.dk)
37. Douglas M, Kroll NM (1974) Ann Phys 82:89
38. Hess BA (1985) Phys Rev A 32:756
39. Hess BA (1986) Phys Rev A 33:3742
40. Benkova Z, Sadlej AJ, Oakes RE, Bell SEJ (2004) J Comput Chem 26:145
41. Peterson KA (2003) J Chem Phys 119:11099
42. Peterson KA, Figgen D, Goll E, Stoll H, Dolg M (2003) J Chem Phys 119:11113
43. Peterson KA, Figgen D, Dolg M, Stoll H (2007) J Chem Phys 126:124101
44. Figgen D, Peterson KA, Dolg M, Stoll H (2009) J Chem Phys 130:164108
45. Shimanouchi T (2001) Molecular vibrational frequencies. In: NIST Chemistry WebBook, No. 69 in NIST Standard Reference Database (National Institute of Standards and Technology, Gaithersburg MD, 20899, 2001), http://webbook.nist.gov
46. Fernández-Sánchez JM, Montero S (1989) J Chem Phys 90:2909
47. Keefe CD, Innis SM (2006) J Mol Struct 785:192