

Mind the basis set superposition error

Demeter Tzeli^{a,*}, Athanassios A. Tsekouras^b

^aTheoretical and Physical Chemistry Institute, National Hellenic Research Foundation, 48 Vassileos Constantinou Ave., Athens GR-11635, Greece

^bLaboratory of Physical Chemistry, Department of Chemistry, National and Kapodistrian University of Athens, Zografou, Athens GR-15771, Greece

ARTICLE INFO

Article history:

Received 17 May 2010

In final form 16 July 2010

Available online 21 July 2010

ABSTRACT

In an ab initio calculation, when the inclusion of core electron correlation is required, the best approach is to use a basis set developed for the electrons that are correlated. However, when a basis set is used that has not been developed for the number of the electrons which are correlated in a calculation, the quantities calculated (geometry, binding energies, frequencies, etc.) need to be corrected for the basis set superposition error. Several cases are discussed and the proper line of action is emphasized.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

In a recent article Hargittai and Vagra [1] stated that ‘quantum chemical calculations have become an everyday tool in chemistry [...] however, many chemists rarely go beyond the routine use of [some commercial] programs, rarely, if ever, checking the original references for basis sets [and such an approach] may lead to unreliable results’. One of the reasons that can lead to unreliable results is failure to correct for the basis set superposition error (BSSE). BSSE is a contribution to the interaction energy arising from the improved description of each fragment in the total basis of the supermolecule compared to the fragment basis alone which results from the inevitable incompleteness of the basis set used in an actual calculation. The BSSE can be corrected by the counterpoise correction (CP) method [2] which is analyzed in Ref. [3]; a detailed example of the procedure for a BSSE calculation can be found in Ref. [4].

Many papers have studied the effect of BSSE, primarily in van der Waals systems and in hydrogen bonded complexes (see for instance [5,6]). It has been shown that the calculation of the BSSE corrections can change their potential energy curves and surfaces [7–10]. Moreover, BSSE has important consequences not only for the bond distances but also for the existence or non-existence of some complexes [7,9–11]. While it is taken for granted that the calculation of the BSSE needs to be taken into account for the above systems [12,13] in the case of a chemically bound molecule with a binding energy of more than a few kcal/mol and a basis set of augmented triple ζ quality or higher, the BSSE should be rather small and, hence, it is ignored [14]. This is true indeed, but under one condition: that the correlated electrons are the same as the electrons for which the basis set has been developed.

It is well known that as we move along the periodic table to heavier elements not only does the correlation of the valence electrons become important, but also the correlation of the inner (core) electrons grows in importance [15]. Moreover, the treatment of the core–valence electron correlation is necessary when high accuracy is needed for energetic and spectroscopic quantities even for molecules involving atoms of the second period of the periodic table [16,17]. Two or three decades ago, it was prohibitive for most theoretical groups to calculate the correlation of the electrons of the inner shells because computational power and storage capacity were limited compared to today. Originally, all the basis sets were developed for correlation of the valence electrons only. The first basis sets modified to account for core–valence electron correlation were published in 1995 [18]. In subsequent years, basis sets developed for core–valence electron correlation were published for more elements [15,19–24].

Thus, it would be expected that when calculations are carried out including the core–valence electron correlation, the appropriate basis set developed for this correlation would be used. However, in many cases the appropriate basis sets are not used, either because they are not available, or because the appropriate basis sets being larger make the calculation more demanding in terms of time and computer resources due to the additional tight core-correlating functions in the basis sets, or because they have not yet been incorporated in some commercial packages.

It has been reported that the use of valence-only basis sets in all electrons calculations can lead to spurious conclusions [25]. It has been shown that the use of the above basis set can produce very unreliable results with a dramatic overestimation of the binding energy and with bond distances which converge in a non-monotonic way to a complete basis set (CBS) limit with increasing basis set size [15,26]. This abnormality is attributed to BSSE [26]. However, in many papers data are obtained from calculations in which all electrons are correlated without taking into account the BSSE. For instance, by searching in the literature of 2009 for MP2(full),

* Corresponding author. Fax: +30 210 7273794.

E-mail address: dtzeli@eie.gr (D. Tzeli).

we found that in 68% of the papers that denoted the use of MP2(full) calculations in their abstract, these calculations had been performed with valence basis sets, for atoms up to the fourth period of the periodic table, without taking into account the necessary BSSE corrections. Thus, a strong statement is needed in the literature to state that this approach is incorrect.

In the present Letter, we report a variety of examples of the BSSE when valence-only basis sets are used and more electrons than the valence electrons are correlated. We point out how BSSE changes with increasing basis set size. We look at diatomic, triatomic and polyatomic molecules, having atoms in the third, fourth and fifth period of the periodic table; their binding energies cover a wide range, from 8 to 200 kcal/mol. All quantities are calculated via ab initio techniques, namely the multireference configuration interaction (MRCI), Møller-Plesset perturbation theory (MP), and coupled cluster (CC) methods. The review of data from the literature is complemented, wherever needed, by additional calculations of our own, specifically on the FeC, All, benzene–Na⁺ and *s*-triazine–Cl[−] systems. The latter calculations were carried out by the GAUSSIAN (MP2) [27] and MOLPRO (RCCSD(T), MRCI) [28] program packages. BSSE corrections were implemented by using the counterpoise correction method [2,4]. The discussion aims at showing the importance of the BSSE and presenting ways in which the results can be improved.

2. Discussion

2.1. Strongly bond molecules

In these cases, the BSSE relative error in binding energy, due to inappropriate valence basis set for the core correlation, may not be as large, but still considerable, if high accuracy is required for the calculated quantities. Moreover, the BSSE can increase with increasing basis set size in contrast to what usually happens when the BSSE is calculated using an appropriate basis set for the number of the correlated electrons. As can be seen in Fig. 1 for the Ga₂N ($\bar{X}^2\Sigma_u^+$) molecule, with a binding energy of 146 kcal/mol with respect to the atomic products [26], the difference between BSSE-corrected and uncorrected Ga–N bond distances, $r_{e\text{-BSSE}}-r_e$, increases from the double to quadruple quality of both augmented and non-augmented valence basis sets. The difference in binding energies $D_e-D_{e\text{-BSSE}}$ increases in a similar way in the case of the augmented valence basis set, but decreases slightly in the case of the non-augmented valence basis set. In contrast, using the appro-

prate correlation consistent-like weighted core basis sets for the number of correlated electrons the BSSE effect in both bond distances and binding energies decreases normally and substantially with the basis set size. Similar diagrams can be derived for the GaN molecule and GaN[−] and Ga₂N[−] anions [26]. The largest relative errors in D_e are 7.8% for Ga₂N ($D_e = 146$ kcal/mol) and 5.8% for Ga₂N[−] ($D_e = 198$ kcal/mol) (at the aug-cc-pVQZ-PP_{Ga}/aug-cc-pVQZ_N level).

Another example of a strongly bound molecule is the FeC ($X^3\Delta$) molecule. Correlating the 4s3d_{Fe}+2s2p_C and the subvalence 3s²3p⁶ electrons, its binding energy, before correcting for BSSE, is 86.4 kcal/mol [29] at the MRCI/aug-cc-pwCV5Z_{Fe}, aug-cc-pV5Z_C level, while its BSSE is only 0.3 kcal/mol. However, at the MRCI/aug-cc-pV5Z_{Fe}, aug-cc-pV5Z_C level (without an appropriate core-valence basis set) the BSSE increases to 1.2 kcal/mol. Finally, with a smaller valence basis set [7s6p4d3f2g1h_{Fe} aug-cc-pVQZ_C] the binding energy is 89.5 kcal/mol [30] with a BSSE of 4.8 kcal/mol. The relative error in the latter case is 5.4%.

Hargittai and Vagra calculated the aluminum monohalides and showed that the use of the valence basis set for the calculation correlating all the electrons may lead to unreliable results and the quadruple basis set quality seems to underestimate the bond length the most [1]. However, a seemingly hopeless situation can be remedied. We examine the All ($X^1\Sigma^+$) molecule as an example. If the BSSE is taken into account, values for the r_e , D_e and ω_e are improved and complete basis set (CBS) limits are established, which, at least for r_e and ω_e , are very close to the experimental values, while the dipole moment appears to converge to a value of 1.95 D (see Table 1 and Fig. 2). The fact that the CBS limit for D_e does not coincide with the experimental value might be due to shortcomings in the experimental data manipulation or in setting up the most appropriate Hamiltonian that describes all the interactions in All. Nonetheless, the existence of a CBS limit shows that for the chosen Hamiltonian the calculation is good.

2.2. Weak or medium strength bonds

In systems with weak or medium strength bonds, the MRCI, CC or MP methods for the calculation of core correlation can lead to BSSE values that are large relative to the binding energy. For example, the binding energy of the Ga₂ molecule, for the correlation of the 4s4p3d electrons, is 31.1 kcal/mol at the RCCSD(T)/aug-cc-pCVQZ-DK level, but increases by 10% when the valence aug-cc-pVQZ-DK basis set is used [15].

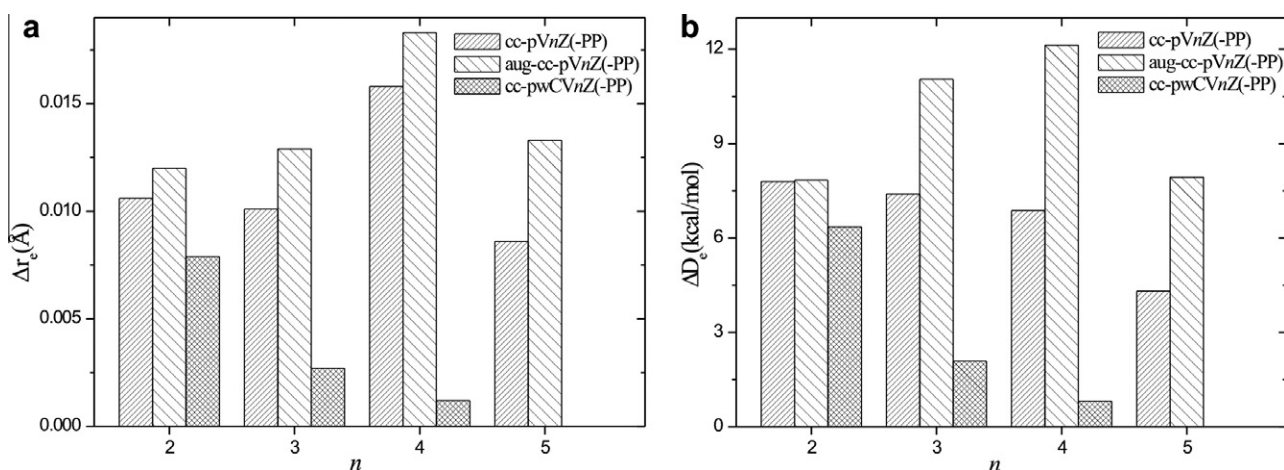


Fig. 1. RCCSD(T) difference in (a) Ga–N bond lengths $\Delta r_e(=r_{e\text{-BSSE}}-r_e)$ in Å and (b) dissociation energy $\Delta D_e(=D_e-D_{e\text{-BSSE}})$ in kcal/mol of Ga₂N ($\bar{X}^2\Sigma_u^+$) with respect to basis set size n , for three types of basis sets, two inappropriate ones, c.f., cc-pVnZ-PP_{Ga}/cc-pVnZ_N and aug-cc-pVnZ-PP_{Ga}/aug-cc-pVnZ_N, and an appropriate one, cc-pwCVnZ-PP_{Ga}/cc-pwCVnZ_N, correlating 31 electrons (4s²4p¹3d¹⁰_{Ga} 2s²2p³_N).

Table 1
Bond lengths R_e (Å), binding energies D_e (kcal/mol), harmonic frequencies ω_e (cm^{-1}) and dipole moments μ (D) and the corresponding BSSE-optimized values of the All ($X^1\Sigma^+$) molecule at the MP2(full)/cc-pVnZ_{AI} cc-pVnZ-PP_I level of theory.

Basis Set	R_e	$R_{e\text{-BSSE}}$	ΔR	D_e	$D_{e\text{-BSSE}}$	ΔD_e	ω_e	$\omega_{e\text{-BSSE}}$	μ	μ_{BSSE}
cc-pVDZ(-PP)	2.574	2.597	0.023	85.19	81.94	3.25	320.5	312.6	1.87	2.12
cc-pVTZ(-PP)	2.539	2.564	0.025	94.29	90.69	3.60	328.1	317.5	1.74	1.94
cc-pVQZ(-PP)	2.499	2.548	0.049	99.24	94.51	4.73	327.2	318.3	1.58	1.94
cc-pV5Z(-PP)	2.518	2.539	0.021	99.70	96.35	3.35	326.8	318.8	1.80	1.95
CBS limit ^d		2.531(2)		101(2)	97.7(1)			318.8(2)		
CCSD(T)/wCQ ^b	2.534 ^c									
MP2(full)/wCT ^d	2.531 ^c			96			321.2			
Expt	2.537 09(3) ^e						316.25(2) ^e			

^a CBS values were estimated using the formula $y = y_{\text{CBS}} + \text{Be}^{-Cn}$ (Ref. [31]).

^b Ref. [1]; CCSD(T)/cc-pwCVQZ_{AI} cc-pwCVQZ-PP_I.

^c The same values are found for the calculation correlating all the electrons and $(n-1)\text{spd} + \text{valence}$ electrons.

^d Ref. [1]; MP2(full)/cc-pwCVTZ_{AI} cc-pwCVTZ-PP_I.

^e Ref. [32].

^f Calculation based on D_0^0 from Ref. [33] and $\omega_e, \omega_e x_e$ from Ref. [32].

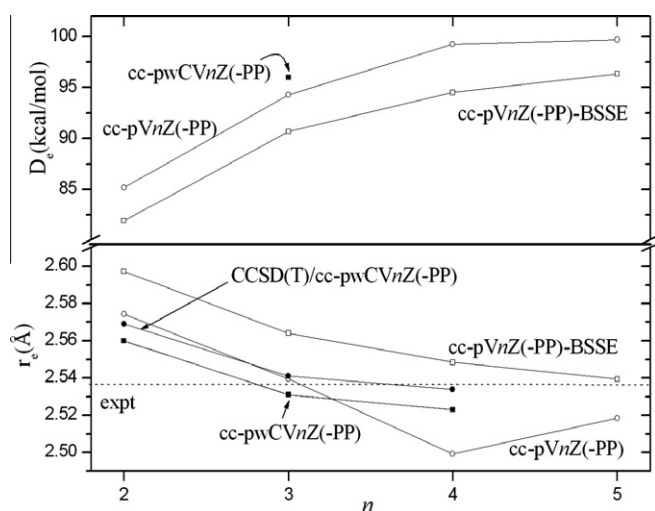


Fig. 2. MP2(full) dissociation energy, D_e , and bond distance, r_e , of the All ($X^1\Sigma^+$) molecule with respect to basis set size n , for cc-pVnZ_{AI}/cc-pVnZ-PP_I, without and with BSSE optimization. Additional data on cc-pVnZ_{AI}/cc-pwCVnZ-PP_I at MP2(full) and RCCSD(T)(full) levels obtained from Ref. [1] are included.

Recently, Estarellas et al. [34] correctly pointed out the erroneous behavior of the MP2(full)/aug-cc-pVnZ, $n = 2, 3$ method in calculations of π -complexes, c.f., benzene–alkaline cations and s -triazine with halide anions. They also concluded that by using the MP2(full)/aug-cc-pCVnZ, $n = 2, 3$ (i.e., a basis set optimized for core–valence electron correlation), they obtained better results. Moreover, in the case s -triazine–Br[−] ($D_{e\text{-BSSE}} = 6.76$ kcal/mol) they found a decrease of 37.5% in its binding energy at the MP2(full)/aug-cc-pVTZ level of theory [34].

An additional conclusion can be extracted from their study regarding the MP2(full)/aug-cc-pVnZ method and the BSSE correction. To illustrate this point, we summarized their pertinent results in Table 2 for benzene–Na⁺ and s -triazine–Cl[−] and filled the gaps with additional calculations of our own. One can see that the MP2(full)/aug-cc-pVTZ method as mentioned by Estarellas et al. [34] has a large BSSE in comparison to MP2/aug-cc-pVTZ and MP2(full)/aug-cc-pCVTZ. However, the optimization taking into account the BSSE correction improves the results significantly and the BSSE-optimized quantities with all three techniques are very similar. Applying the BSSE correction modifies the shape of the potential energy surface. This is clear in the case of benzene–Na⁺ at the MP2(full)/aug-cc-pVTZ level, where R_e is elongated by 0.2 Å when the BSSE is taken into account showing its large effect.

In general, the large BSSE arises because the basis set developed to account for valence electron correlation is poorer and less complete when used for the correlation of core–valence electrons than for the valence electrons only. The effective quality of a basis set used in this way is degraded and the nominal quality of the basis set can be misleading, primarily for elements beyond the second period. As required by the variation theorem, the absolute energies of the elements in the third period including the correlation of the ($2s^2 2p^6$) electrons are higher when the basis set is not modified for the core–valence electron correlation. For example, absolute energies are higher at the RCCSD(T)/aug-cc-pVQZ≡[7s6p4d3f2g] level than the values at the RCCSD(T)/aug-cc-p(w)CVDZ≡[6s5p3d] level of theory (see Table 3), i.e., the use of valence quadruple quality basis set yields worse results than the core–valence double zeta basis set. This also happens for the elements of the fourth period of periodic table [26]. However, if the BSSE is taken into account, the geometries and the dissociation energies of the molecules studied are significantly improved. In the case of the Ga_{*n*}N^{*x*} species, $n = 1, 2$ and $x = 0, -1$, it has been shown that the CBS limits of binding energies and bond distances with core–valence electron correlation are practically the same for both core–valence and valence basis sets when the BSSE is incorporated [26].

In some cases it appears that for the double and maybe for the triple quality of valence basis set, the uncorrected quantities r_e and D_e are better than the BSSE-corrected values (see for example Ga_{*n*}N^{*x*} [26], AlX [1]), yet this is only due to cancellation of errors. An inappropriate basis set yields a small binding energy and a large BSSE; correction for the latter then yields an even smaller binding energy. However, this does not happen, in the case of the quadruple valence basis set, and the BSSE-corrected results are much better. Thus, the proper approach is at least to correct the BSSE

Table 2

Interaction energies without (ΔE), with (ΔE_{BSSE}), and optimized with ($\Delta E_{\text{BSSE-opt}}$) BSSE correction in kcal/mol and bond distances of π -system–ion, without (R_e) and with BSSE optimization ($R_{e\text{-BSSE}}$) in Å for benzene–Na⁺ and s -triazine–Cl[−] complexes as given in Ref. [34].

Method	ΔE	ΔE_{BSSE}	$\Delta E_{\text{BSSE-opt}}$	R_e	$R_{e\text{-BSSE}}$
Benzene–Na ⁺					
MP2/aug-cc-pVTZ	−22.56	−22.00	−22.00 ^a	2.468	2.475 ^a
MP2(full)/aug-cc-pVTZ	−31.08	−21.18	−22.62	2.235	2.445
MP2(full)/aug-cc-pCVTZ	−25.23	−24.01	−24.07	2.384	2.405
<i>s</i> -Triazine–Cl [−]					
MP2/aug-cc-pVTZ	−8.53	−7.76	−7.78 ^a	3.143	3.181 ^a
MP2(full)/aug-cc-pVTZ	−9.69	−7.81	−7.86	3.119	3.187
MP2(full)/aug-cc-pCVTZ	−8.67	−7.83	−7.85	3.141	3.181

^a Additional calculations of this study.

Table 3

RCCSD(T) absolute energies (hartree) of Na, Si and Cl, correlating the 2s2p3s3p electrons.

Basis Set	Na	Si	Cl
aug-cc-pVQZ	−161.871057	−288.980816	−459.755332
aug-cc-p(w)CVDZ	−162.044123 ^a	−289.059297 ^b	−459.766414 ^b
aug-cc-p(w)CVTZ	−162.130955 ^a	−289.193621 ^b	−459.954364 ^b
aug-cc-p(w)CVQZ	−162.163074 ^a	−289.244723 ^b	−460.021306 ^b

^a aug-cc-pCVnZ.

^b aug-cc-pwCVnZ.

because the results corrected for BSSE correspond to the basis set which is used.

3. Conclusion

To summarize, when the inclusion of core electron correlation is required, the best approach is to use the post HF method combined with a basis set developed for the electrons that are correlated. If this is not possible, the quality of the valence basis set used is effectively reduced, especially for elements beyond the second period of the periodic table. The situation can be amended by correcting for the BSSE. Potential energy surface or curve calculation and structure optimization must be carried out while taking into account the BSSE with the counterpoise method, a procedure which yields significantly better results. Finally, it should be noted that with an inappropriate basis set the BSSE can increase with increasing basis set size in contrast to what usually happens when the BSSE is calculated using an appropriate basis set for the number of the correlated electrons.

Acknowledgments

Financial support from the EU FP7, Capacities Program, NANO-HOST project (GA 201 729) is acknowledged.

References

- [1] M. Hargittai, Z. Varga, *J. Phys. Chem.* 111 (2007) 6.
- [2] S.F. Boys, F. Bernardi, *Mol. Phys.* 19 (1970) 553.
- [3] F.B. van Duijneveldt, J.G.C.M. van Duijneveldt-van de Rijdt, J.H. van Lenthe, *Chem. Rev.* 94 (1994) 1873.
- [4] S.S. Xantheas, *J. Chem. Phys.* 104 (1996) 8821.
- [5] A. Halkier, W. Klopper, T. Helgaker, P. Jørgensen, P.R. Taylor, *J. Chem. Phys.* 111 (1999) 9157 (and references therein).
- [6] T. van Mourik, A.K. Wilson, K.A. Peterson, D.E. Woon, T.H. Dunning, *Adv. Quantum Chem.* 32 (1999) 105.
- [7] T. van Mourik, *J. Phys. Chem. A* 112 (2008) 11017.
- [8] D. Tzeli, I.D. Petsalakis, G. Theodorakopoulos, *J. Phys. Chem. A* 111 (2007) 8892.
- [9] A.E. Shields, T. van Mourik, *J. Phys. Chem. A* 111 (2007) 13272.
- [10] L.F. Holroyd, T. van Mourik, *Chem. Phys. Lett.* 442 (2007) 42.
- [11] R. Crespo-Otero, L.A. Montero, W.-D. Stöhrer, J.M.G. De la Vega, *J. Chem. Phys.* 123 (2005) 134107.
- [12] D. Tzeli, A. Mavridis, S.S. Xantheas, *J. Chem. Phys.* 112 (2000) 6178.
- [13] D. Tzeli, A. Mavridis, S.S. Xantheas, *J. Phys. Chem. A* 106 (2002) 11327.
- [14] T.H. Dunning Jr., *J. Phys. Chem. A* 104 (2000) 9062.
- [15] N. DeYonker, K.A. Peterson, A.K. Wilson, *J. Phys. Chem. A* 111 (2007) 11383.
- [16] K.A. Peterson, T.H. Dunning Jr., *J. Chem. Phys.* 106 (1997) 4119.
- [17] K.A. Peterson, A.K. Wilson, D.E. Woon, T.H. Dunning Jr., *Theor. Chem. Acc.* 97 (1997) 251.
- [18] D.E. Woon, T.H. Dunning Jr., *J. Chem. Phys.* 103 (1995) 4572.
- [19] K.A. Peterson, C. Puzzarini, *Theor. Chem. Acc.* 114 (2005) 283.
- [20] K.A. Peterson, D. Figgen, M. Dolg, H. Stoll, *J. Chem. Phys.* 126 (2007) 124101.
- [21] N.B. Balabanov, K.A. Peterson, *J. Chem. Phys.* 123 (2005) 064107.
- [22] K.A. Peterson, T.H. Dunning Jr., *J. Chem. Phys.* 117 (2002) 10548.
- [23] C. Hättig, *Phys. Chem. Chem. Phys.* 7 (2005) 59.
- [24] J.G. Hill, S. Mazumder, K.A. Peterson, *J. Chem. Phys.* 132 (2010) 4572.
- [25] D. Feller, K.A. Peterson, *Chem. Phys. Lett.* 430 (2006) 459 (and references therein).
- [26] D. Tzeli, A.A. Tsekouras, *J. Chem. Phys.* 128 (2008) 144103.
- [27] M.J. Frisch et al., GAUSSIAN 09, Revision A.1.
- [28] H.J. Werner, P.J. Knowles, et al., MOLPRO 2006.1 is a Package of ab initio Programs.
- [29] D. Tzeli, A. Mavridis, *J. Chem. Phys.* 132 (2010) 194312.
- [30] D. Tzeli, A. Mavridis, *J. Chem. Phys.* 118 (2003) 4984.
- [31] D. Feller, *J. Chem. Phys.* 96 (1992) 6104.
- [32] F.C. Wyse, W. Gordy, *J. Chem. Phys.* 56 (1972) 2130.
- [33] E. Martin, R.F. Barrow, *Phys. Scr.* 17 (1978) 501.
- [34] C. Estarellas, X. Lucas, A. Frontera, D. Quiñonero, P.M. Deyà, *Chem. Phys. Lett.* 489 (2010) 254.