

Counterpoise corrected interaction energies are not systematically better than uncorrected ones: comparison with CCSD(T) CBS extrapolated values

Juan Raul Alvarez-Idaboy · Annia Galano

Received: 23 October 2009 / Accepted: 26 October 2009 / Published online: 13 November 2009
© Springer-Verlag 2009

Abstract The effect of the inclusion of counterpoise corrections (CP) on the accuracy of interaction energies has been studied for different systems accounting for (1) intermolecular interactions, (2) intramolecular interactions and (3) chemical reactions. To minimize the error associated with the method of choice, the energy calculations were performed using CCSD(T) in all the cases. The values obtained using aug-cc-pVXZ basis sets are compared to CBS-extrapolated values. It has been concluded that at least for the tested systems CP corrections systematically leads to results that differ from the CBS-extrapolated ones to a larger extension than the uncorrected ones. Accordingly, from a practical point of view, we do not recommend the inclusion of such corrections in the calculation of interaction energies, except for CBS extrapolations. The best way of dealing with basis set superposition error (BSSE) is not to use CP corrections, but to make a computational effort for increasing the basis set. This approach does not eliminate BSSE but significantly decreases it, and more importantly it proportionally decreases all the errors arising from the basis set truncation.

Keywords Basis set superposition error (BSSE) · Counterpoise correction · Weak interactions · Reaction barriers

J. R. Alvarez-Idaboy (✉)
Departamento de Física y Química Teórica,
Facultad de Química, Universidad Nacional Autónoma
de México, 04510 Mexico DF, Mexico
e-mail: jidaboy@unam.mx

A. Galano (✉)
Departamento de Química, Universidad Autónoma
Metropolitana-Iztapalapa, 09340 Mexico DF, Mexico
e-mail: agalano@prodigy.net.mx

1 Introduction

The term basis set superposition error (BSSE) was introduced by Liu and McLean in 1973 [1], although it was first reported in 1969 by Kestner [2] for the helium dimer and by Jansen and Ros [3] for the protonation of carbon monoxide. In the last decades it has been extensively studied and reviewed [4–19]. However, it remains a controversial subject. BSSE arises when two (or more) fragments approach to form a supermolecule and it is a consequence of the basis set incompleteness. The description of fragment *A* within the complex is improved by the basis functions of fragment *B* and vice versa, while such extension is not possible in the calculation of the isolated fragments. Therefore, the energy of the fragments in the complex ($E_{A\{AB\}}, E_{B\{AB\}}$) becomes lower than in the isolated fragments (E_A, E_B). Since the interaction energy associated with the supermolecule formation (E_{IAB}) is calculated as:

$$E_{IAB} = E_{AB}^{\text{Opt}} - E_A^{\text{Opt}} - E_B^{\text{Opt}} \quad (1)$$

where the superscript Opt denotes that their geometries have been optimized, the energy release involved in the *AB* formation decreases due to two factors: the actual stabilization of the system due to the interaction of the fragments and the artificial improvement of their description. This second effect (BSSE) is a computational artifact, which causes an unphysical overestimation of the interaction energy.

It is important to notice that BSSE is independent of the kind of interaction involved in the formation of the supermolecule and that for a given basis set, it is expected to increase as the fragments become closer. It is often considered when studying the interaction of intermolecular complexes, but only at that point of the potential energy surface (PES). Recently, Kobko and Dannenberg [19] have pointed out the importance of

taking BSSE into account when computing transition states (TS) where the same problem is expected. However, when an addition reaction is studied and the energy associated with the product formation is computed usually the BSSE is not taken into account. This is a surprising fact since the magnitude of the BSSE is expected to increase in the following order: encounter complex < TS < addition product. This is the logical order based on the very nature of the BSSE and on the proximity of the fragments at each point of the PES. BSSE is present in any molecular quantum chemical calculation using the LCAO approach in conjunction with finite Gaussian basis sets, and it should be taken into account for all energy calculations relative to isolated fragments or for none of them. It is inconsistent to consider BSSE only for weakly bonded complexes. In fact the importance of BSSE in strongly bound systems has been previously pointed out [20, 21].

To remove the artificial energy lowering caused by the BSSE a positive correction ($\delta_{AB}^{\text{BSSE}}$) needs to be included in Eq. (1):

$$EI_{AB} = E_{AB}^{\text{Opt}} - E_A^{\text{Opt}} - E_B^{\text{Opt}} + \delta_{AB}^{\text{BSSE}} \quad (2)$$

The most widely used approach to correct BSSE is the counterpoise correction (CP^n , n = number of fragments) proposed by Boys and Bernardi [22, 23]. When the energy of each fragment is calculated with the full basis set of the supermolecule (but without the nuclei or electrons of the other fragments), using so-called “ghost orbitals”, this approach is referred to as full-counterpoise (FCP). For a two-fragment scheme, Eq. (1) becomes:

$$EI_{AB}^{\text{FCP}^2} = E_{AB} - E_{A\{AB\}} - E_{B\{AB\}} \quad (3)$$

where the energies of the fragments are re-computed with the $\{AB\}$ basis sets, but are optimized using the fragment geometries without the ghost orbitals. In this approach the BSSE correction becomes:

$$\delta_{AB}^{\text{BSSE}} = (E_A + E_B) - (E_{A\{AB\}} + E_{B\{AB\}}) \quad (4)$$

Another source of error within the CP scheme is the need of geometry re-optimization within the CP approach. Some researchers re-optimize the geometry with the counterpoise correction included, and in some cases we will also explore that option here. The equation for doing this is

$$EI_{AB}^{\text{FCP}^2} = E_{AB}^{\text{Opt}} - E_{A\{AB\}}^{\text{Opt}} - E_{B\{AB\}}^{\text{Opt}} \quad (5)$$

where the geometries of the fragments are re-optimized with the $\{AB\}$ basis sets.

Over the years the use of FCP approach has become a controversial topic. It has been proposed that it overcorrects the BSSE (some examples can be found in 4, 5, 7–12,

and references therein). One argument supporting this overcorrection is based on the fact that, the basis set that is actually available to a fragment in the supermolecule is not the complete basis set of its partner but only the space of the unoccupied orbitals. Accordingly, the interaction energies, calculated using Eq. (3), become less attractive than they should be. A different approach, known as virtual counterpoise (VCP), was proposed by Daudey et al. [5]. In the VCP scheme the energies of the fragments in the supermolecule are calculated using a basis that contains the ghost virtual orbital space but not the ghost occupied orbitals. However, van Duijneveldt et al. [14] argued that the full CP scheme is conceptually correct and does not overcorrect BSSE.

The CP approach is also affected by the number of fragments chosen to represent the interaction. If the two-fragment scheme is used, problems of definition arise for open shell species interacting with neutral molecules. For this particular problem, a three-fragment approach has also been tested [16–19]. However, the CP correction values increase with the number of fragments, increasing the uncertainty on the reliability of this kind of corrections. Other ambiguity issues, within the two-fragment CP schemes, have also been reported. For the HF trimer, for instance, different BSSE corrections are obtained depending on the definition of the fragments, i.e. the correction to the H...F interaction calculated for HF...((HF...HF)) is different than that for ((HF...HF))...HF [24].

Recently, a new CP approach that eliminates the fragment-definition issues has been proposed by us [25]. It is known as “atom-by-atom” scheme (CP^{aa}). Within the CP^{aa} scheme the intramolecular BSSE of all the involved species are computed considering every atom as a fragment. Then the BSSE correction associated with the supermolecule formation is obtained by subtracting the intramolecular BSSE of the fragments from the BSSE of the supermolecule:

$$\delta_{AB}^{\text{BSSE}}(CP^{\text{aa}}) = CP_{AB}^{\text{aa}} - CP_A^{\text{aa}} - CP_B^{\text{aa}} \quad (6)$$

This approach not only eliminates the fragment-definition problems but also significantly decreases the reported counterpoise overcorrection of the A–B interaction, even though it includes all the ghost orbitals, and not only the unoccupied ones.

In addition to the issues discussed above, about CP corrections of BSSE, it is important to notice that the corrections are intended to “improve” the interaction energy between two fragments, i.e. to make theoretical descriptions better reproducing the actual phenomena. However, BSSE is not the only error arising from the basis set incompleteness. At Hartree–Fock level the truncation of the basis set introduces errors in the total energy of any modeled system. Accordingly, these errors are present in

the calculation of the energies of A , B and AB , and necessarily lead to an error in the interaction energy at any level of theory. However, the sign of this error is not necessarily known. Additionally, when A and B approach the electrons of A correlate their movement with the electrons of B , which causes the London dispersion forces. These forces are attractive but with truncated basis sets, and limited configuration interaction methods, these interaction energies are underestimated. By analogy with BSSE we are naming this error basis set correlation error (BSCE). BSCE and BSSE have opposite sign. However, while correcting BSSE has become a common practice, correction for the BSCE is not included into the interaction energy estimations. Accordingly, the BSSE “overcorrection” also arises from correcting only one of two errors of opposite sign. Accordingly, it seems worthwhile to wonder if including BSSE corrections is the best approach for practical purposes, i.e. for reproducing the actual behavior of the studied systems. At this point it seems that the controversy about BSSE has overcome the CP overcorrection issue and discussion should be focused on whether the inclusion of BSSE corrections improves or doesn't improve the interaction energies.

From a pragmatic point of view it has been reported that sometimes the use of CP correction does not lead to improved interaction energies [6, 7]. Therefore it has been proposed not to use CP corrections, but instead to make a computational effort for increasing the basis set. This approach does not eliminate BSSE but significantly decreases it, and most importantly it proportionally decreases all the errors arising from the basis set truncation. Surprisingly, this theoretically sound suggestion has been generally ignored and the CP correction is extensively used. At the time these suggestions were made it was not possible to get reliable BSSE-free energies and to prove that the CP corrections do not improve the interaction energies with respect to its magnitude at infinite basis set. Nowadays it is possible to do that using extrapolations to complete basis set (CBS) for relatively small systems and comparing these results, which are BSSE-free, with those obtained by “correcting” with the CP scheme. Direct comparisons with experiment are not an alternative procedure since the calculated energies are always affected by the intrinsic errors of the chosen method and not only by the incompleteness of the basis set.

In a very interesting paper by Dunning [26] he stated: “It is quite possible and even probable that the binding energies computed without the counterpoise correction, are closer to the complete basis set limit than the corrected values. This frustrating (or lucky) situation (depending on your proclivities) is due to the fact that the BSSE and the basis set convergence error are often of opposite signs”. On

the other hand the CP scheme is widely used with the purpose of increasing the accuracy of calculations. Therefore, the main goal of this work is to test if the inclusion of CP corrections actually improves the computed data, for commonly used basis sets, taking CBS approaches as reference. Obviously this test must be performed on a limited set of systems, thus they were selected to include interactions of different nature.

2 Computational details

All the calculations were performed with the Gaussian 03 packages of programs [27]. Different systems in which BSSE could be important have been considered. They correspond to three different kinds: (1) intermolecular interactions, (2) intramolecular interactions and (3) chemical reactions. The dimers of argon, HF and HCl have been chosen as representative of the first kind. In the Ar dimer the attractive interaction is purely due to correlation energy and the choice of the number of fragments is unambiguous. Moreover, these systems are of special importance because the interaction energies are very small and therefore the geometry is very sensitive to the level of the calculation. Even though both HF and HCl dimers are formed through hydrogen bonding, they are different in nature. While HF dimer constitutes a classical hydrogen-bonded system, HCl dimer is a system bound mainly by dipole forces. The F_2 molecule has been chosen for studying intramolecular interactions. In addition the H abstraction from formaldehyde by OH radical and the OH addition to acetylene have been modeled as examples of chemical reactions.

To minimize the error associated with the method of choice, the energy calculations were performed using CCSD(T) in all the cases. Geometry optimizations were performed at the same level when possible. When such procedure became computationally unfeasible, the geometries have been obtained using other strategies that are clarified in the section corresponding to each particular case. For systems with more than two atoms we have tested conventional and atom-by-atom CP approaches. Within the CP^{aa} approach the atomic fragments were modeled in their ground state, i.e. H, F, Cl: charge 0, multiplicity 2; C, O: charge 0, multiplicity 3; Ar: charge 0, multiplicity 1.

Extrapolations to complete basis set, with and without CP corrections, have been performed using the aug-cc-pVXZ ($X = 2, 3, 4, 5$) or ($X = 2, 3, 4$) correlation consistent basis sets [28]. The extrapolations were performed using the Klopper type [29] expression:

$$f(X) = f_{\text{CBS}} + \frac{a}{X^3} + \frac{b}{X^4} \quad (7)$$

Theoretically both ways of CBS extrapolation (with and without including CP corrections) should converge to the same value. Any disagreement between them could be attributed to several problems, such as extrapolation procedure, chosen method, basis set quality, etc. This subject has been recently addressed by Varandas [30], and it was found that extrapolation without CP is an excellent choice. In the present work both extrapolations are reported but all the errors are calculated with respect to the CP-corrected extrapolated values.

3 Results and discussion

3.1 The Ar dimer

In this case we have used aug-cc-pV(X+d)Z basis sets, since they have been proven to better describe molecules containing second row atoms [31]. The binding energies ($BE = E_A + E_B - E_{AB}$) with (BE^{CP}) and without (BE) CP corrections are reported in Table 1, for the Ar dimer. The deviations of BE^{CP} and BE from the value obtained by CBS (BE^{CBS}) extrapolation are also reported in this table.

It was found that the general trend of the uncorrected binding energies is to increase with the basis set size. It was found that the CP-corrected BE values are systematically worse than the uncorrected ones. Even for the smallest tested basis set the deviation of the BE uncorrected value is 0.059 kcal/mol, while the deviation of the corrected value is 0.157 kcal/mol. The same trend was obtained for all the other tested basis sets. Accordingly,

Table 1 Binding energies with and without CP corrections, and their deviations from the value obtained by CBS extrapolation (all in kcal/mol) for argon dimer

	BE	BE^{CP}	$BE^{CBS} - BE$	$BE^{CBS} - BE^{CP}$
Using uncorrected geometries				
aug-cc-pV(D+d)Z	0.223	0.124	0.059	0.158
aug-cc-pV(T+d)Z	0.272	0.200	0.010	0.082
aug-cc-pV(Q+d)Z	0.271	0.237	0.011	0.045
aug-cc-pV(5+d)Z	0.285	0.259	-0.003	0.023
aug-cc-pV(6+d)Z	0.278	0.268	0.004	0.014
CBS-extrapol.	0.280	0.282		
Using CP-corrected geometries				
aug-cc-pV(D+d)Z		0.125		0.157
aug-cc-pV(T+d)Z		0.201		0.081
aug-cc-pV(Q+d)Z		0.237		0.045
aug-cc-pV(5+d)Z		0.259		0.023
aug-cc-pV(6+d)Z		0.268		0.014
CBS-extrapol.		0.282		

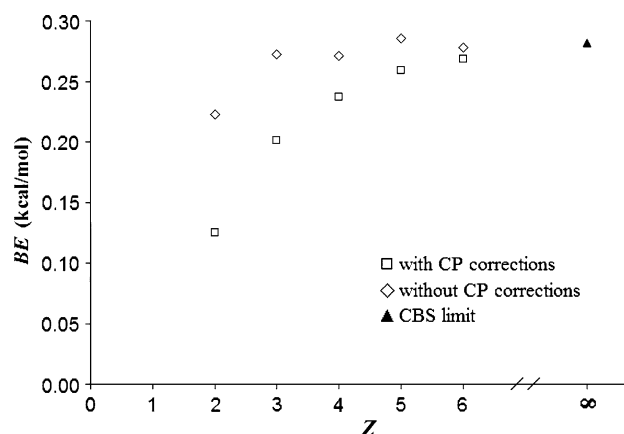


Fig. 1 Plot of binding energies (BE , in kcal/mol) obtained with and without CP-corrections with different basis sets, compared to that obtained at CBS limit

correcting only the BSSE would certainly worsen the BE values compared to CBS limit. However, as the basis set increases the error of the CP-corrected values decreases, while the error of the uncorrected BE values shows a slight zig-zag behavior. This leads to a smoother convergence for the former ones. The uncorrected CBS extrapolated and the CP-corrected CBS-extrapolated values found for the Ar_2 system are identical, as theoretically expected, and in agreement with the previous report by Varandas for similar system [30]. It should be noticed that the zig-zag pattern for the uncorrected values arises from the very small errors from the aug-cc-pV(T+d)Z basis set up. In fact the BE error using this basis set is smaller than the BE^{CP} error when using aug-cc-pV(6+d)Z, i.e. there is almost a perfect cancellation of errors between BSSE and BSCE.

The trends on the binding energies obtained with and without CP-corrections with different basis sets are shown in Fig. 1. In this figure it becomes evident that as the size of the basis set increases the differences in the values of the BE computed with and without CP-corrections vanishes, and that they converge to the CBS limit. The figure also shows that non-corrected values are closer to the CBS limit for all the tested basis sets, despite the fact that the convergence of the uncorrected values is less smooth (zig-zag pattern).

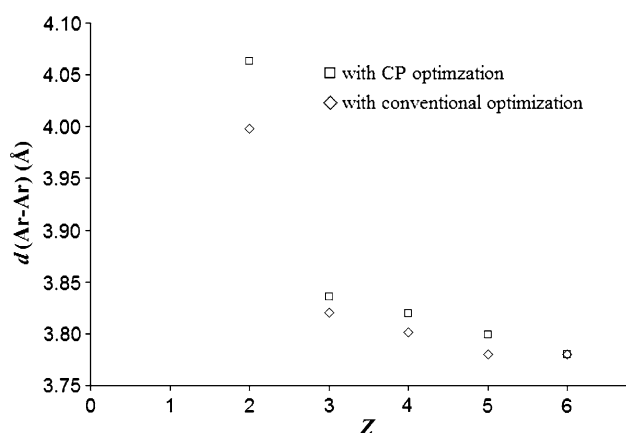
Geometry optimizations of the Ar dimer were performed with and without CP-corrected gradients (Table 2). It was found that the Ar–Ar distance in the dimer decreases with the increase of the basis set with both methodologies but converging to the same value (3.780 Å). For basis sets from aug-cc-pV(D+d)Z to aug-cc-pV(5+d)Z the Ar–Ar distance is larger when CP correction is used, which is the expected behavior. The difference between CP-corrected and uncorrected distances decreases with the size of the

Table 2 Ar–Ar distance (Å) obtained from optimizations with and without CP-corrected gradients

	$d(\text{Ar}-\text{Ar})$	$d(\text{Ar}-\text{Ar})^{\text{CP}}$
aug-cc-pV(D+d)Z	3.998	4.063
aug-cc-pV(T+d)Z	3.821	3.836
aug-cc-pV(Q+d)Z	3.801	3.820
aug-cc-pV(5+d)Z	3.780	3.799
aug-cc-pV(6+d)Z	3.780	3.780

basis set, with a largest difference of 0.065 Å for the aug-cc-pV(D+d)Z. The dependence of the bond distance with the size of basis set was found to be less marked for the conventional optimizations. The Ar–Ar distance was found to vary by 0.283 and 0.218 Å (aug-cc-pV(D+d)Z to aug-cc-pV(6+d)Z) for CP-corrected and uncorrected geometries, respectively. The binding energies obtained when using CP-corrected geometries are very similar to those obtained from conventional geometry optimizations (Table 1). However, the $\text{BE}^{\text{CBS}} - \text{BE}^{\text{CP}}$ difference slightly decreases, as expected, when the geometries are optimized using the CP-corrected gradients.

The trend on the Ar–Ar distance as the size of the basis set increases, when the optimizations are carried out using CP-corrected gradients is compared with that obtained from conventional geometry optimizations in Fig. 2. This figure clearly shows that the larger difference in the reported geometrical parameter occurs when the smallest basis set is used, and that for aug-cc-pV(6+d)Z they become identical. Accepting that the converged geometry is the correct one, the conventional optimization always leads to better geometrical description of the Ar–Ar dimer.

**Fig. 2** Plot of Ar–Ar distance (Å) obtained from optimizations with and without CP-corrected gradients with different basis sets

Since aug-cc-pV(D+d)Z is a relatively small basis set, the most widely accepted approach would probably be to include CP corrections in the binding energy calculations. Therefore, perhaps the most important comparison, in terms of binding energies, is that between the CP-corrected values for this basis set and the CBS-extrapolated value. According to values in Table 1, the uncorrected binding energy is underestimated by 0.059 kcal/mol, which represents a 21% error with respect to the CBS-extrapolated energy, for non-CP corrected geometry. Since the uncorrected binding energy is already underestimated, the CP corrections would actually worsen the result, leading to an underestimation of 0.158 kcal/mol (56% error). This large error might be attributed to the use of Ar₂ uncorrected geometry. However, when the dimer geometry is CP-corrected the error scarcely diminishes, even though the geometry change is significant. Accordingly, it can be stated that this basis set is not large enough to properly describe such weak interaction and it should not be used for that purpose. The drastic truncation of the basis set certainly cannot be corrected by including CP corrections in the calculation of the binding energies.

The results up to this point might be enough to demonstrate that the use of CP corrections do not represent a systematic improvement in binding energies calculations, since it is not possible to know a priori if such corrections would actually improve the results. However, a similar analysis needs to be performed for other systems in order to elucidate if this case is the exception or the rule.

3.2 The F₂ molecule

The BSSE in a chemically bonded molecule is not a well-studied phenomenon, but as it was pointed out previously the BSSE is present in any system with more than one atom [20, 21]. This is the reason for which it is recommended to use CP correction in chemical reaction when bonds are breaking or forming [19]. Therefore, there is no reason to ignore it when energies of reaction are calculated. Kobko and Dannenberg [19] proposed to use a three-fragment CP correction for methane + CH₃ reaction. Within this scheme the intramolecular BSSE of one of the C–H bonds of methane was calculated.

Since the atoms are closer when involved in chemical bonds than they are in intermolecular complexes, the magnitude of the BSSE is even larger in the former case. However, its relative importance is minor because the energy evolution associated with the chemical reaction is usually significantly larger than the BSSE. We have used the F₂ molecule as an example for illustrating the importance of the intramolecular BSSE.

The values in Table 3 show that the trends are very similar than those previously discussed for the Ar dimer, in

Table 3 Binding energies with and without CP corrections, and their deviations from the value obtained by CBS extrapolation (all in kcal/mol) for the F₂ molecule

	BE	BE ^{CP}	BE ^{CBS} – BE	BE ^{CBS} – BE ^{CP}
Using uncorrected geometries				
aug-cc-pVDZ	29.859	27.254	8.438	11.042
aug-cc-pVTZ	36.388	34.856	1.909	3.441
aug-cc-pVQZ	37.601	36.811	0.696	1.486
aug-cc-pV5Z	37.991	37.768	0.306	0.529
CBS-extrapol.	38.297	38.586		
Using CP-corrected geometries				
aug-cc-pVDZ		27.268	–	11.029
aug-cc-pVTZ		34.864	–	3.433
aug-cc-pVQZ		36.812	–	1.485
aug-cc-pV5Z		37.768	–	0.529
		38.297		

spite of a dramatic change in the nature of the interaction and in the magnitude of the binding energy. The CBS-extrapolated value was found to be equal to 38.586 kcal/mol. Increasing the basis set size provokes a systematic increase in the binding energy from 29.859 kcal/mol (aug-cc-pVDZ) to 37.991 kcal/mol (aug-cc-pV5Z), when CP corrections are not included. There is a large discrepancy between the DZ result and the CBS-extrapolated one (8.727 kcal/mol). Therefore, since by definition the CP correction diminishes the binding energy, its inclusion will worsen the value of the interaction energy regardless of the size of the basis set. Clearly, the largest error is obtained when the smallest basis set is used. The discrepancy between BE^{CP} and BE^{CBS} is 11.332 kcal/mol, when using DZ basis set, i.e. about 2.6 kcal/mol larger than that calculated without including CP corrections.

The behavior of the BE compared to the BE^{CP} is very similar when using CP-corrected geometries, with differences no larger than 0.014 kcal/mol. The values of the F–F distance corresponding to conventional geometry optimizations and using the CP-corrected gradients are shown in Table 4. With the exception of the values obtained with aug-cc-pVDZ set, all the *d*(F–F) values are very similar. In fact they can be considered identical from aug-cc-pVQZ

Table 4 F–F distance (Å) obtained from optimizations with and without CP-corrected gradients

	<i>d</i> (F–F)	<i>d</i> (F–F) ^{CP}
aug-cc-pVDZ	1.450	1.457
aug-cc-pVTZ	1.418	1.423
aug-cc-pVQZ	1.413	1.415
aug-cc-pV5Z	1.411	1.411

regardless of the inclusion of the CP-corrected gradients. For the smallest of the tested basis sets (DZ), the discrepancy from the converged value was found to be slightly larger if the geometry is optimized within the CP frame. Accordingly it represents a useless increase in computational cost to optimize the geometry using CP-corrected gradients, even for the smallest of the tested basis sets. It should be noticed that the only difference with the Ar dimer, discussed above, is that the convergence of the interatomic distance is much faster for the F₂ molecule, which is a logical finding due to the different nature of both interactions.

3.3 The HF dimer

To keep the level of theory as high as possible and to minimize the computational cost the geometrical parameters for this system were taken from a previous publication by Peterson and Dunning [32]. In that work the geometries were fully optimized at CCSD(T)/aug-cc-pVTZ while for aug-cc-pVQZ partial optimizations were performed for the F–F distance. The other bond distances were extrapolated and the aug-cc-pVTZ bond angles were used. In the present work we have used CBS extrapolations DZ–5Z instead of DZ–QZ, and the geometrical parameters for aug-cc-pV5Z are those obtained for aug-cc-pVQZ. In this case geometries were not optimized using CP correction, since it makes the computational cost almost prohibiting with no significant gain in the quality of the energy results, as demonstrated above. In this case the supermolecule contains more than two atoms, i.e. the fragments themselves are polyatomic systems. Accordingly, in addition to the conventional two-fragment CP approach (CP²), the atom-by-atom corrections (CP^{aa}) were also used to account for BSSE. For this and all the following systems the CP-corrected CBS-extrapolated values used as reference for computing BE errors are those obtained from the CP^{aa} approach. However, if the errors were computed with respect to the conventional CP extrapolations the conclusions would be the same.

The binding energies for the HF dimer do not present the systematic behavior of the Ar₂ and F₂ systems (Table 5). The nature of the interaction is more complex, since hydrogen bond and dipole–dipole interactions are present. Another difference with Ar₂ and F₂ systems is that the attractive interaction in the HF dimer involves an H atom.

A total BSSE of 1.73 kcal/mol was obtained for the intramolecular BSSE in the HF molecule using aug-cc-pVTZ basis set. This value represents the sum of the energy improvement of the F atom when computed including the ghost orbitals of the H atom ($E_{F-F} - E_{F(HF)}$) and the equivalent improvement in the description of the H atom ($E_{H-F} - E_{H(HF)}$):

Table 5 Binding energies without CP corrections and with CP² and CP^{aa} corrections, and their deviations from the value obtained by CBS extrapolation (all in kcal/mol) for the HF dimer

	BE	BE ^{CP2}	BE ^{CPaa}	BE ^{CBS} – BE	BE ^{CBS} – BE ^{CP2}	BE ^{CBS} – BE ^{CPaa}
aug-cc-pVDZ	4.813	4.023	4.163	–0.232	0.558	0.417
aug-cc-pVTZ	4.829	4.313	4.377	–0.248	0.268	0.204
aug-cc-pVQZ	4.724	4.486	4.512	–0.143	0.095	0.069
aug-cc-pV5Z	4.652	4.525	4.547	–0.071	0.056	0.034
CBS-extrapol.	4.581	4.593	4.581			

$$\delta_{HF}^{BSSE} = \delta_F^{BSSE} + \delta_H^{BSSE} \quad (8)$$

where:

$$\delta_F^{BSSE} = E_F - E_{F\{HF\}} \quad (9)$$

$$\delta_H^{BSSE} = E_H - E_{H\{HF\}} \quad (10)$$

It was found that $\delta_F^{BSSE} = 1.7$ kcal/mol and $\delta_H^{BSSE} = 0.03$ kcal/mol. Therefore, the BSSE due to the incompleteness of the basis set for the H atom can be considered negligible (1.76% of the total BSSE). As a consequence, in interactions where the key role is played by H atoms, one of the fragments is less affected by BSSE and the total BSSE is smaller than that for those cases where the intermolecular interaction does not involve H atoms. For this reason intermolecular associations ruled by interactions involving H atoms are not the best example for generalizations about the proper use of CP corrections. Unfortunately due to the importance of hydrogen bonds, these interactions have been widely used as examples in the study of BSSE. In fact the most widely studied system, related to BSSE, is the water dimer.

The HF dimer is also a special case in the sense that a Hartree–Fock calculation of the binding energy with a relatively small basis set (aug-cc-pVDZ) gives a binding energy of 3.85 kcal/mol with errors of 0.75 and 0.73 kcal/mol compared to the experimental and CCSD(T) extrapolated values, respectively. The only explanation for such good agreement is a fortuitous cancellation of errors. Using CP² correction the HF/aug-cc-pVDZ binding energy becomes equal to 3.67 kcal/mol. i.e. the BSSE error is 0.18 kcal/mol. Increasing the basis set to aug-cc-pVQZ does not change the CP²-corrected binding energy, in spite of the fact that the BSSE error diminishes to 0.05 kcal/mol which could be attributed to the decrease of the dipole moment from 1.899 to 1.884. It also can be attributed to the decrease of the electron sharing effects or the intermolecular delocalization when the basis set is increased. To test that we have performed NBO calculations of the HF dimer using Hartree–Fock and M05-2X functional with

cc-pVDZ and cc-pV5z basis sets. The results were found to be qualitatively independent of the used method and show that the charge transfer is larger when using the smaller basis set: 0.020 and 0.012 electrons for DZ and 5Z basis sets, respectively, and the M05-2X functional. The E(2) energy, associated with the intermolecular delocalization is larger (12.01 kcal/mol) using the DZ basis set than with the 5Z basis set (8.93 kcal/mol). This means that the covalent component of the binding energy of the HF dimer decreases as the basis set increases. In other words, this a case where the sign of the error due to the basis set truncation have the same sign as the BSSE and opposite to the error due to the incomplete CI which is larger for the smallest basis set. This denotes that including BSSE could actually improve or at least not worsen the energies for this particular case and probably for conventional H bonds. However, this behavior cannot be extrapolated to other systems involving a different kind of interaction. In other words, these results are valid for H-bonded systems, and H-bonded systems only.

Values in Table 5 show that going from aug-cc-pVDZ to aug-cc-pVTZ the uncorrected BE increases, while it decreases from aug-cc-pVTZ to aug-cc-pVQZ. This is a different trend than those previously described for Ar dimer and F₂ molecule. Comparing uncorrected and CP-corrected binding energies with the CBS-extrapolated value it can be noticed that the uncorrected BE are overestimated and the BE^{CP2} and BE^{CPaa} values are underestimated. For the aug-cc-pVDZ basis set the largest absolute value of the discrepancy was obtained within the BE^{CP2} scheme, while the uncorrected value shows the best agreement with the CBS extrapolation. This is a clear example that the generalized conception that the smaller the basis set the largest the improvement by including BSSE is not always fulfilled, at least when using the conventional CP² approach. For the aug-cc-pVTZ basis set the error of the uncorrected value is of similar magnitude but opposite sign than the CP-corrected ones. For the other two basis sets the largest discrepancy with the CBS-extrapolated value arises from uncorrected values,

but the BSSE is almost negligible. Comparing the results obtained by including CP² and CP^{aa} corrections, it can be noticed that the latter are slightly closer to the CBS-extrapolated ones.

3.4 The HCl dimer

Geometry optimizations were performed for this system at the CCSD level using analytical gradients in conjunction with TZ basis sets.

In contrast with the HF dimer, which is mainly formed through electrostatic interactions and partial covalent bonding, dispersion interactions are more important in the formation of the HCl dimer. Accordingly it seems interesting to compare the binding energy results for both systems. It was found that, as in the case of the HF dimer, the uncorrected BE are overestimated and the BE^{CP2} and BE^{CPaa} values are underestimated, regardless of the used basis set. The underestimation from BE^{CPaa} approach was found to be lower than that from BE^{CP2} scheme. In this case when the aug-cc-pVDZ basis set is used the largest absolute value of the discrepancy with the CBS extrapolation was also obtained within the BE^{CP2} approach, while it arises from uncorrected values for the rest of the tested basis sets (Table 6).

Despite the difference in the relative importance of the various kinds of forces leading to the formation of the HF and HCl dimers, both systems behave similarly in terms of the BSSE influence on the binding energy. In both cases the interaction of the two fragments involves an H atom, in contrast to the Ar₂ and F₂ examples.

3.5 The formaldehyde + OH reaction

For this system geometry optimizations were also performed at the CCSD level using analytical gradients in conjunction with TZ basis sets. Due to the size of the system the calculations were performed only up to aug-cc-pVQZ. In this case we will focus on the BSSE influence on the barrier of reaction ($\Delta E = E_{\text{TS}} - E_{\text{Reac}}$). This is of particular interest since ΔE is the main parameter when comparing the viability of different reaction paths, and also represents the energetic feature that rules kinetic calculations.

This reaction is known to occur exclusively by H abstraction and with small negative activation energy [33]. Accordingly the transition state structure can be thought as a three-fragment system: HC(O)⋯H⋯OH. Thus, the CP corrections have been performed within the CP³ and CP^{aa} schemes. The most significant finding in this case is that regardless of the used basis set the CP-corrected barriers differ from the CBS extrapolation to a larger extent than the uncorrected ones, with smaller differences for CP^{aa} than for CP³ scheme (Table 7). Even the uncorrected ΔE obtained with the smallest of the tested basis set (aug-cc-pVDZ) is closer to the extrapolated one than those obtained using the largest (aug-cc-pVQZ) and counterpoise corrections. Accordingly, the use of CP corrections worsens the interaction energy (ΔE), i.e. from a practical point of view is a better approach to use the uncorrected values. Actually the largest discrepancy with ΔE^{CBS} from the uncorrected calculations is -0.407 kcal/mol, while the smallest discrepancy from the CP³ and CP^{aa}-corrected barriers are -0.678 and -0.485 kcal/mol, respectively.

Table 6 Binding energies without CP corrections and with CP² and CP^{aa} corrections, and their deviations from the value obtained by CBS extrapolation (all in kcal/mol) for the HCl dimer

	BE	BE ^{CP2}	BE ^{CPaa}	BE ^{CBS} – BE	BE ^{CBS} – BE ^{CP2}	BE ^{CBS} – BE ^{CPaa}
aug-cc-pVDZ	2.258	1.537	1.567	–0.280	0.441	0.411
aug-cc-pVTZ	2.152	1.781	1.828	–0.174	0.197	0.150
aug-cc-pVQZ	2.049	1.889	1.909	–0.071	0.089	0.069
aug-cc-pV5Z	2.008	–	–	–0.030	–	–
CBS-extrapol.	1.952	1.996	1.978			

Table 7 Electronic energy barriers without CP corrections and with CP³ and CP^{aa} corrections, and their deviations from the value obtained by CBS extrapolation (all in kcal/mol) for the formaldehyde + OH reaction

	ΔE	ΔE^{CP3}	ΔE^{CPaa}	$\Delta E^{\text{CBS}} - \Delta E$	$\Delta E^{\text{CBS}} - \Delta E^{\text{CP3}}$	$\Delta E^{\text{CBS}} - \Delta E^{\text{CPaa}}$
aug-cc-pVDZ	–0.352	1.27	0.487	–0.407	–2.029	–1.246
aug-cc-pVTZ	–0.427	0.349	0.128	–0.332	–1.108	–0.887
aug-cc-pVQZ	–0.468	–0.081	–0.274	–0.291	–0.678	–0.485
CBS-extrapol.	–0.665	–0.527	–0.759			

Table 8 BSSE (in kcal/mol) for electronic barriers and energies of reaction calculated according to CP² and CP^{aa} schemes for acetylene + OH reaction

	Barrier		Energy of reaction	
	CP	CP ^{aa}	CP	CP ^{aa}
aug-cc-pVDZ	2.076	1.345	4.944	2.872
aug-cc-pVTZ	0.781	0.474	1.940	1.149
aug-cc-pVQZ	0.315	0.169	0.780	0.445

3.6 The acetylene + OH reaction

In this case geometry optimizations were also performed at the CCSD level using analytical gradients in conjunction with TZ basis sets and, as for the previous system, the calculations were performed only up to aug-cc-pVQZ due to the size of the system. In the particular case of addition reactions it seems interesting to analyze BSSE not only for the barrier but also for the energy of reaction, since it is expected to increase as the fragments become closer. If this assumption is correct, BSSE should be larger for the latter interaction energy since the fragments are as close as possible in the addition product, and according to the results in Table 8 that is exactly the case. Therefore, if BSSE corrections are going to be included in the computation of a specific reaction, there is no reason to do that in the calculation of the interaction energies of the weak complexes and not in the transition state where the fragments are closer. Furthermore, if the studied reaction represents an addition process, then the BSSE should be included along the whole reaction path. This would be the only coherent way to include BSSE corrections.

The values corresponding to reaction barriers and energies of reaction calculated within BSSE uncorrected, CP² and CP^{aa} schemes are shown in Tables 9 and 10, respectively.

The best agreement with the CBS-extrapolated values was found for those calculations that do not include any CP correction, regardless of the used basis set. With the exception of the uncorrected result, obtained with DZ basis set, all the other calculated barriers are higher than the CBS-extrapolated value. Since the inclusion of CP

corrections to BSSE also increases the interaction energies, such inclusions invariably worsen the results. Similar effect is found for the energies of reaction, since they are all less negative than the CBS-extrapolated ones.

3.7 Related studies

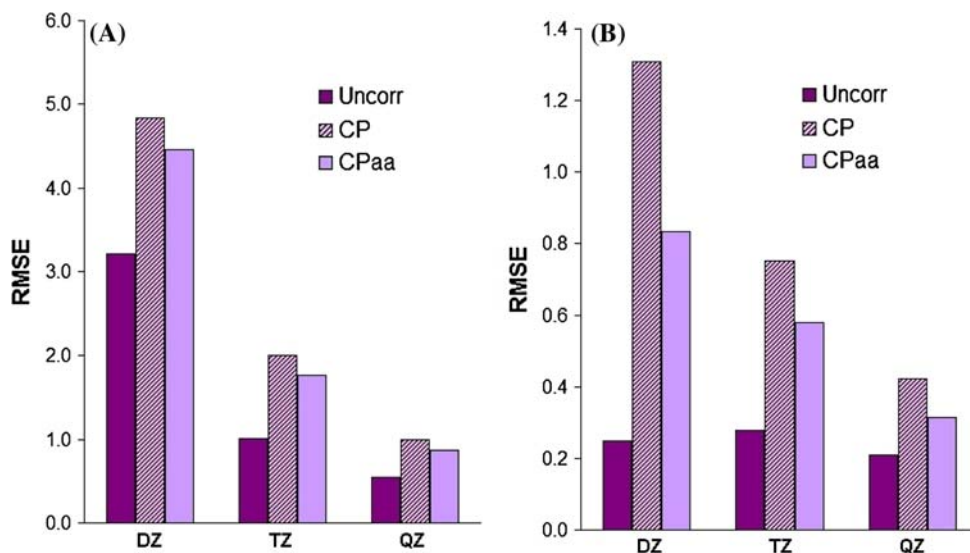
There are several publications that support the results presented above. Gruber-Stadler et al. [34] have found that even using the CP^{aa} approach at RHF-RCCSD(T)/aug-cc-pVTZ level of theory, the corrected activation energy of Cl + formaldehyde reaction is overestimated compared to the CBS-extrapolated value. Accordingly, they concluded that the error due to the truncation of the basis set is larger than the BSSE, and for that reason they excluded the correction from the kinetic calculations. In a recent work from our group [35] we arrive to similar conclusion in the study of isoprene and butadiene + OH reactions. The CBS-QB3 and uncorrected MP2/6-311++G(d,p) results are in excellent agreement with experimental rate coefficient. However, the rate coefficient calculated using CP-corrected MP2/6-311++G(d,p) energies disagrees with the experimental value in several orders of magnitude which corresponds to an overestimation of the activation energy by almost 5 kcal/mol. Papajak et al. [36] have shown that increasing the size of the basis set does not necessarily decrease the CP correction. In particular they have demonstrated (averaged over 18 cases) that smaller basis sets can produce smaller counterpoise corrections than larger basis sets. Dabkowska et al. [37] studied H-bonded and stacked uracil dimers and found that the uncorrected geometries with a small basis set 6-31g(d) are equivalent to the counterpoise corrected cc-pVTZ ones. That finding also supports the conclusions from the present work. Moreover, from our experience calculating many atmospheric reactions at high level of theory [38 and references 27, 28, 87, 122, 140, 169–171 therein; 39–44] we always get an excellent agreement with experiment when using uncorrected energies, while any attempt to use CP-corrected energies worsen the agreement with experiments. This may be the reason why in computational kinetics the use of counterpoise corrections is exceptional.

Table 9 Electronic energy barriers without CP corrections and with CP² and CP^{aa} corrections, and their deviations from the value obtained by CBS extrapolation (all in kcal/mol) for the acetylene + OH reaction

	ΔE	ΔE^{CP^2}	$\Delta E^{\text{CP}^{\text{aa}}}$	$\Delta E^{\text{CBS}} - \Delta E$	$\Delta E^{\text{CBS}} - \Delta E^{\text{CP}^3}$	$\Delta E^{\text{CBS}} - \Delta E^{\text{CP}^{\text{aa}}}$
aug-cc-pVDZ	0.267	2.343	1.612	0.097	-1.979	-1.248
aug-cc-pVTZ	0.795	1.576	1.268	-0.431	-1.212	-0.904
aug-cc-pVQZ	0.692	1.006	0.86	-0.328	-0.642	-0.496
CBS-extrapol.	0.471	0.358	0.364			

Table 10 Electronic energies of reaction without CP corrections and with CP² and CP^{aa} corrections, and their deviations from the value obtained by CBS extrapolation (all in kcal/mol) for the acetylene + OH reaction

	ΔE	ΔE^{CP^2}	$\Delta E^{\text{CP}^{\text{aa}}}$	$\Delta E^{\text{CBS}} - \Delta E$	$\Delta E^{\text{CBS}} - \Delta E^{\text{CP}^3}$	$\Delta E^{\text{CBS}} - \Delta E^{\text{CP}^{\text{aa}}}$
aug-cc-pVDZ	-33.985	-29.041	-31.113	-0.861	-5.805	-3.733
aug-cc-pVTZ	-33.093	-31.153	-31.944	-1.753	-3.693	-2.902
aug-cc-pVQZ	-33.685	-32.905	-33.24	-1.161	-1.941	-1.606
CBS-extrapol.	-34.607	-34.927	-34.846			

Fig. 3 Root mean square errors (RMSE) with respect to CBS-extrapolated values for: **a** all the studied systems, **b** the subset of systems involving only weak interactions

4 Concluding remarks

It is clear that using only six systems, even if they are carefully selected to take into account as many kind of interactions as possible, we cannot conclude that CP corrections always worsen the interaction energies. However, it is enough to prove that CP does not always improve interaction energies as frequently assumed.

To help to visualize the results from the present work altogether, root mean square errors (RMSE) have been calculated with respect to the CBS-extrapolated values. The results are shown in Fig. 3, where they have been displayed in two different ways: including all the studied systems (A), and including only the subset of systems that involves weak interactions (B), i.e. excluding the F₂ system and the energy of reaction for the acetylene + OH reaction. It is clear from Fig. 3 that the discrepancies with CBS-extrapolated values are smaller for the CP^{aa} scheme than for the conventional two (or three)-fragment counterpoise correction. This finding also supports the hypothesis that the conventional CP schemes overcorrect for BSSE, compared to CP^{aa}.

The most relevant information from Fig. 3 is that the inclusion of CP corrections systematically leads to results that differ from the CBS-extrapolated ones to a larger extent

than the uncorrected ones. Contrary to the general belief, this effect is more significant if only weak interactions are analyzed. Accordingly, from a practical point of view, when relatively small basis sets are used (DZ and TZ), we do not recommend the inclusion of such corrections in the calculation of interaction energies since it would lead to values with larger discrepancies with the accurate ones. Accordingly, the best way of dealing with BSSE, for large-sized systems that make CBS extrapolations computationally unfeasible, is not to use CP corrections, but instead to make a computational effort for increasing the basis set. This approach does not eliminate BSSE but significantly decreases it, and more importantly it proportionally decreases all the errors arising from the basis set truncation.

Acknowledgments The authors would like to thank the Dirección General de Servicios de Cómputo Académico (DGSCA) at Universidad Nacional Autónoma de México and Laboratorio de Visualización y Computo Paralelo at UAM-Iztapalapa. This work was partially supported by a grant from the DGAPA UNAM (PAPIIT-IN203808).

References

1. Liu B, Malean AD (1973) *J Chem Phys* 59:4557
2. Kestner NR (1968) *J Chem Phys* 48:252

3. Jansen HB, Ros P (1969) *Chem Phys Lett* 3:140
4. Johansson A, Kollman P, Rothenberg S (1973) *Theor Chim Acta* 29:167
5. Daudey JP, Claverie P, Malrieu P (1974) *Int J Quantum Chem* 8:1
6. Frisch MJ, Del Bene JE, Binkley JS, Schaefer HF III (1986) *J Chem Phys* 84:2279
7. Schwenke DW, Truhlar DG (1985) *J Chem Phys* 82:2418
8. Morokuma K, Kitaura K (1981) JHG. In: Politzer P (ed) *Chemical application of atomic and molecular electronic potentials*. Plenum, New York
9. Lopez JC, Alonso JL, Lorenzo FJ, Rayon VM, Sordo JA (1999) *J Chem Phys* 111:6363
10. Hunt SW, Leopold KR (2001) *J Phys Chem A* 105:5498
11. Valdés H, Sordo JA (2002) *J Comput Chem* 23:444
12. Valdés H, Sordo JA (2002) *J Phys Chem A* 106:3690
13. Chalasinski G, Szczesniak MM (1994) *Chem Rev* 94:1723
14. van Duijneveldt FB, de Rijdt JGCM, van Lenthe JH (1994) *Chem Rev* 94:1873
15. Kestner NR, Combarías JE (1999) *Rev Comp Chem* 13:99
16. Wieczorek R, Haskamp L, Dannenberg JJ (2004) *J Phys Chem A* 108:6713
17. Ponti A, Mella M (2003) *J Phys Chem A* 107:7589
18. Salvador P, Duran M, Dannenberg JJ (2002) *J Phys Chem A* 106:6883
19. Kobko N, Dannenberg JJ (2001) *J Phys Chem A* 105:1944
20. Wilson AK, van Mourik T, Dunning TH (1996) *J Mol Struct: Theochem* 388:339
21. van Mourik T, Wilson AK, Peterson KA, Woon DE, Dunning TH (1998) *Adv Quantum Chem* 31:105
22. Boys SF, Bernardi F (1970) *Mol Phys* 19:553
23. Meunier A, Levy B, Berthier G (1973) *Theor Chim Acta* 29:49
24. Turi L, Dannenberg JJ (1993) *J Phys Chem* 97:2488
25. Galano A, Alvarez-Idaboy JR (2006) *J Comput Chem* 27:1203
26. Dunning TH Jr (2000) *J Phys Chem A* 104:9062
27. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, 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 JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) *Gaussian 03, Revision D.01*. Gaussian Inc., Wallingford
28. Woon DE, Dunning TH (1993) *J Chem Phys* 98:1358
29. Klopffer W, Bak KL, Jørgensen P, Olsen J, Helgaker T (1999) *J Phys B At Mol Opt Phys* 32:R103
30. Varandas AJC (2008) *Theor Chem Acc* 119:511
31. Dunning TH Jr, Peterson KA, Wilson AK (2001) *J Chem Phys* 114:9244
32. Peterson KA, Dunning TH Jr (1995) *J Chem Phys* 102:2032
33. Alvarez-Idaboy JR, Mora-Diez N, Boyd RJ, Vivier-Bunge A (2001) *J Am Chem Soc* 123:2018
34. Gruber-Stadler M, Muhlhauser M, Sellevag SR, Nielsen CJ (2008) *J Phys Chem A* 112:9
35. Francisco-Márquez M, Alvarez-Idaboy JR, Galano A, Vivier-Bunge A (2008) *Chem Phys* 344:273
36. Papajak E, Leverentz HR, Zheng J, Truhlar DG (2009) *J Chem Theory Comput* 5:1197
37. Dabkowska I, Jurecka P, Hobza P (2005) *J Chem Phys* 122:204322
38. Galano A, Alvarez-Idaboy JR (2008) In: Goodsite ME, Johnson MS (eds) *Advances in quantum chemistry: applications of quantum chemistry to the atmosphere*, vol 55, chap 12. Elsevier, Amsterdam, p 245
39. Alvarez-Idaboy JR, Reyes L, Mora-Diez N (2007) *Org Biomol Chem* 5:3682
40. Alvarez-Idaboy JR, Galano A, Bravo-Perez G, Ruiz ME (2001) *J Am Chem Soc* 123:8387
41. Alvarez-Idaboy JR, Mora-Diez N, Vivier-Bunge A (2000) *J Am Chem Soc* 122:3715
42. Vega-Rodriguez A, Alvarez-Idaboy JR (2009) *Phys Chem Chem Phys* 11:7649
43. Francisco-Márquez M, Alvarez-Idaboy JR, Galano A, Vivier-Bunge A (2003) *Phys Chem Chem Phys* 7:1392
44. Uc VH, Alvarez-Idaboy JR, Galano A, Vivier-Bunge A (2008) *J Phys Chem A* 112:7608