Structural manifestation of the delocalization error of density functional approximations: C_{4N+2} rings and C_{20} bowl, cage, and ring isomers

Tim Heaton-Burgess and Weitao Yang[a)]
Department of Chemistry, Duke University, Durham, North Carolina 27708, USA

(Received 1 December 2009; accepted 13 May 2010; published online 18 June 2010)

The ground state structure of C_{4N+2} rings is believed to exhibit a geometric transition from angle alternation ($N \leq 2$) to bond alternation ($N > 2$). All previous density functional theory (DFT) studies on these molecules have failed to reproduce this behavior by predicting either that the transition occurs at too large a ring size, or that the transition leads to a higher symmetry cumulene. Employing the recently proposed perspective of delocalization error within DFT we rationalize this failure of common density functional approximations (DFAs) and present calculations with the rCAM-B3LYP exchange-correlation functional that show an angle-to-bond-alternation transition between C_{10} and C_{14}. The behavior exemplified here manifests itself more generally as the well known tendency of DFAs to bias toward delocalized electron distributions as favored by Hückel aromaticity, of which the C_{4N+2} rings provide a quintessential example. Additional examples are the relative energies of the C_{20} bowl, cage, and ring isomers; we show that the results from functionals with minimal delocalization error are in good agreement with CCSD(T) results, in contrast to other commonly used DFAs. An unbiased DFT treatment of electron delocalization is a key for reliable prediction of relative stability and hence the structures of complex molecules where many structure stabilization mecahnisms exist. © *2010 American Institute of Physics*. [doi:10.1063/1.3445266]

I. INTRODUCTION

Density functional theory (DFT) is a popular approach for molecular modeling and in particular studying large carbon-based systems where competing effects such as Hückel aromaticity, electron correlations, and symmetry breaking are significant and must be approached with rigor while maintaining computationally feasibility. With linear scaling approaches well developed the opportunity to investigate ever more complex systems with DFT is an attractive prospect. Yet, there are many situations where density functional approximations (DFAs) to the exchange-correlation functional of DFT can generate unacceptable errors.

One well known issue of DFT that is of particular importance to large carbon structures is the tendency of popular DFAs to bias toward delocalized electron distributions or fractional charges.[1,2] This observation has long been noted in small molecules where DFAs often overbind systems where delocalized electron distributions are a possibility.[3] Such biasing can significantly distort molecular geometries as has been seen clearly in oligomers, where bond length alternation (BLA) is suppressed by all common forms of non-long-range corrected DFAs.[4,5] In its most extreme form this failure of DFAs can completely negate stabilizing mechanisms giving rise to incorrect molecular geometries as will be discussed here with C_{4N+2} rings. In exhibiting the behavior of these rings for various DFAs we demonstrate an effect of the delocalization error (DE)—a concept defined through the behavior of systems with noninteger total number of electrons. That this concept has such consequences as radically distorting equilibrium geometries, and increasing HOMO-LUMO gaps, is important for understanding when commonly employed density functional will provide poor results, and encouraging the development of functionals that address, among other issues, the DE. Irrespective of the presence of delocalization mechanisms, DFA thermochemistries tend to be poorer as system size increases.[6] Addressing thermochemistry through constraint satisfaction and fitting is the most popular approach in recent DFA development.[7–9] Unlike many other constraints, the effect of mistreatment of the DE is well known and widely studied.

Four ground state structures are considered for these rings, as shown in Fig. 1: a high symmetry cumulene ($D_{(4N+2)h}$—all bond lengths and bond angles are equal) together with two reduced symmetry ($D_{(2N+1)h}$) forms—an angle-alternating cumulene and a bond length-alternating (acetylenic) structure, and a further symmetry reduced species ($C_{(2N+1)h}$) with both bond and angle alternation. These structures provide a sensitive test for the unbiased treatment of delocalization by DFAs owing to the presence of a second order Jahn–Teller effect stabilizing the reduced symmetry forms relative to the high symmetry cumulene. So far, DFAs have been unable to describe correctly the structural evolution of this system. We will show that this failure of existing exchange-correlating functionals can be avoided with a correct understanding of the delocalization properties of DFAs, and as such marks the first application of a DFA designed to treat delocalization in an unbiased manner to this system which has a history of conflicting results when comparing results obtained with different DFAs.[10–15] The behavior we observe here, and more importantly the fundamental understanding that has been recently provided which allows us to

a)Electronic mail: weitao.yang@duke.edu.

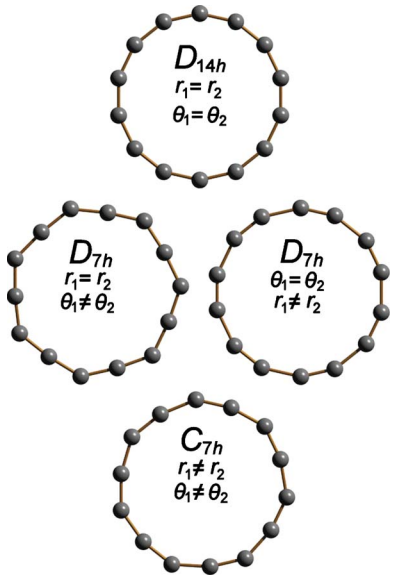

FIG. 1. Conceivable ground state structures for the C_{4N+2} rings as exemplified by C_{14}.

correctly deal with these systems, will be essential for the future development of exchange-correlation functionals suitable for dealing with large carbon structures where many stabilizing mechanisms compete directly with Hückel aromaticity. We also present calculations on the bowl, cage, and ring isomers of C_{20}, another system with a confusing history,[16] where the presence of a DE may affect the energetic ordering.

Many of the systematic failures of DFAs are due to the DE[1,2] of common DFAs. These include the overestimation of oligomer polarizabilities[17] and the underestimation of band gaps[18] and reaction barriers.[19] The DE is the incorrect behavior of the ground state energy as a function of the number of electrons. The exact many-body ground state energy is a piecewise linear function interpolating the energy at integral numbers of electrons.[20,21] Deviation from this straight line behavior represents an error—a convex curve for the total energy, as is typical of most DFAs, stabilizes the energy at noninteger numbers of electrons and is termed the DE. Conversely, a concave curve, as obtained from Hartree–Fock (HF) theory, represents a localization error. The DE is most naturally observed in the fragmentation of molecules to infinite separation, where it leads to electrons delocalized over each fragment and too low a total energy. The relationship between delocalization and localization errors in stabilizing delocalized/localized density distributions at near-equilibrium geometries is less transparent. Perhaps the most prominent structural effect that has been correlated with the DE is the BLA observed in oligomers,[4] where the alternation predicted by common DFAs is too small (the electron distribution is too delocalized), and decays too rapidly with the oligomer length, compared to wave function theories. Long-range corrected functionals such as CAM-B3LYP and LC-PBE have been shown to significantly improve the description of BLA relative to hybrids, GGAs and LDA.[5] This trend follows that of the DE typically seen in each class of functional.[22]

The highest levels of calculations that have been performed on the C_{4N+2} system are the quantum Monte Carlo (QMC) and limited coupled cluster (CCSD) calculations of Torelli and Mitas.[10] These calculations predicted a crossover from angle- to bond-alternating structures between C_{10} and C_{14}, with the higher symmetry cumulene present as a transition state. It should be noted that the structures employed by Torelli and Mitas are approximations to the local minima as they were obtained by a sampling force under bond length variations rather than a full geometry optimization.

HF calculations show the same transition as QMC, albeit significantly overestimating the relative stability of the bond alternating form which can be rationalized as a manifestation of the localization error of HF. DFT calculations on this system yielded strongly functional-dependent results, all of which conflict with the QMC and HF results. LDA and GGA functionals show a transition from angle alternation to the higher symmetry cumulene between C_{10} and C_{14}. These functionals have the largest DE resulting in an overstabilization of the aromatic structure to the extent of completely negating the second order Jahn–Teller effect. Only extrapolating to the infinite limit has given rise to a stable bond-alternating structure as induced by a Peierls mechanism.[23] The B3LYP hybrid functional improves upon this result,[12,13] giving rise to the correct transition from angle to bond alternation but this occurs between C_{18} and C_{22} (or between C_{22} and C_{26}).[14] The inclusion of HF exchange decreases the DE and we see this manifested in the better performance of B3LYP. A recent CCSD study by Arulmozhiraja and Ohno[15] predicted a structure with both bond and angle alternation to be the most stable form for C_{14}, C_{18}, and C_{22}, and the bond-alternating structure to be more stable than the angle-alternating cumulene. Yousaf and Taylor[24] have studies of some small C_{4N} and C_{4N+2} clusters, demonstrating that single reference coupled cluster and multireference SCF struggle to describe the ground state structures, and that B3LP performs somewhat similarly to single reference coupled cluster. Watts and Bartlett[25] similarly considered detailed coupled cluster and many-body perturbation theory on the C_{10} ring, suggesting the angle-alternating structure to be the most stable.

In the following we argue that this failure of popular DFAs is a result of the DE fully (with LDA, GGAs) or partially negating the second order Jahn–Teller effect, and show that when the DE is taken into account by the exchange-correlation functional correct results are obtained.

II. RESULTS

Only recently have exchange-correlation functionals been explicitly designed in an attempt to minimize the DE. Two functionals, MCY3 and rCAM-B3LYP, attempt to do so by including a fit to the straight line behavior of the total energy for the ionization and electron capture of a carbon atom when constructing the functionals.[22] Here we employ rCAM-B3LYP, a reparametrization of CAM-B3LYP,[26] implemented within a modified version of the DALTON code[27] to perform geometry optimizations on C_{4N+2} rings with the 6-311G* basis set. Calculations with the CAM-B3LYP,

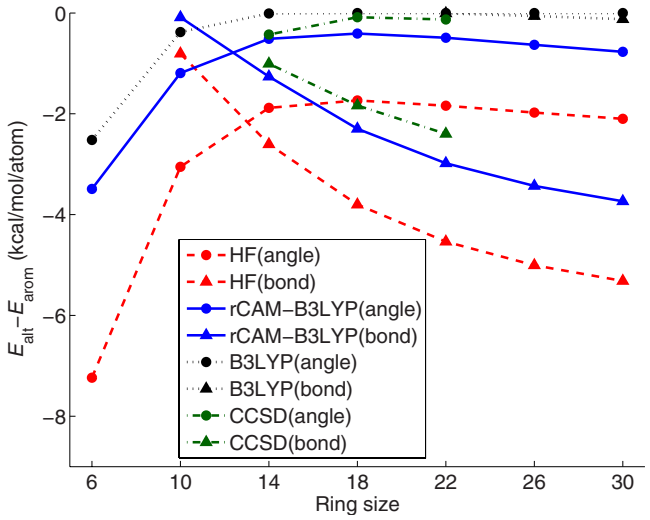

FIG. 2. Energies of the bond- and angle-alternating structures relative to the aromatic ($D_{(4N+2)h}$) structure as a function of the ring size. rCAM-B3LYP gives rise to the structural transition between C_{10} and C_{14}. The CCSD values are taken from Arulmozhiraja and Ohno (Ref. 15) (employing the 6-31G* basis set).

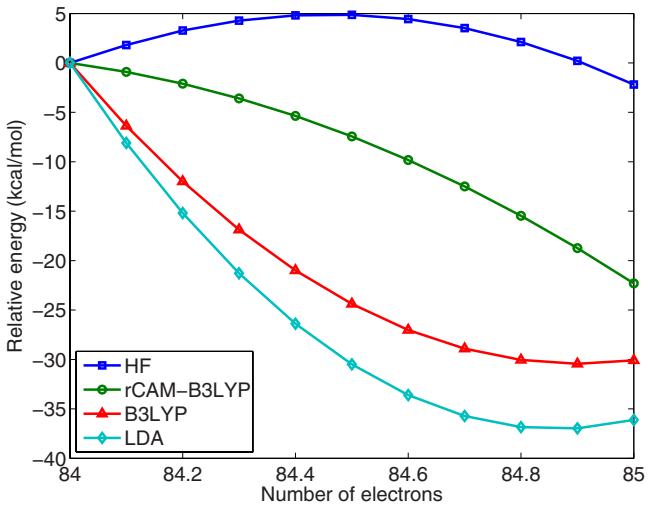

FIG. 3. Relative energy of C_{14} (most stable structure of each functional) with fractional number of electrons.

B3LYP, PBE0, PBE, BLYP, and LDA (SVWN5) functionals, and with HF theory, were also performed.

Figure 2 displays the energies of the angle- and bond-alternating structures relative to the aromatic form calculated with the rCAM-B3LYP and B3LYP functionals and within HF and CCSD (Ref. 15) theory. For clarity we exclude those functionals from Fig. 2 for which the bond-alternating structure is not stable for small rings (LDA, BLYP, and PBE) and the CAM-B3LYP results. The observed behavior of all the functionals studied is summarized in Table I.

The rCAM-B3LYP functional predicts a transition from angle to bond alternation to occur between C_{10} and C_{14} in agreement with the QMC calculations of Torelli and Mitas,[10] albeit over stabilizing the alternating structures with respect to the QMC results (by ~10 kcal/mol for C_{10} and C_{14}). The CAM-B3LYP functional, which was not designed with the DE in mind but which displays a better fractional charge behavior than hybrid functionals, predicts a transition from angle-to-angle and bond alternation between C_{10} and C_{14}, in agreement with the CCSD calculations of Arulmozhiraja and Ohno.

The CCSD results of Arulmozhiraja and Ohno predict a bond- and angle-alternating structure to be most stable. We have been unable to converge to such a structure with rCAM-B3LYP and the 6-311G* basis set. However, when using 6-31G*, D95(d,p), or cc-pVDZ basis sets that Arulmozhiraja and Ohno employed, we do observe this structure to be the lowest in energy indicating some basis set effects with our rCAM-B3LYP DFT approach. Calculations employing larger basis sets also failed to converge to such a structure. Comparing rCAM-B3LYP relative energies (ignoring the bond- and angle-alternating structure) to these CCSD results we see a much better agreement than with QMC. With the 6-31G* basis set, CCSD predicts the C_{14} acetylenic and cumulenic forms to lay 14.05 and 5.98 kcal/mol below the aromatic structure, respectively, with rCAM-B3LYP yielding 17.64 and 7.56 kcal/mol. The BLA, the difference in bond lengths divided by the mean bond length, is shown in Fig. 3 for various functionals and the CCSD and QMC results discussed above. For C_{14}, rCAM-B3LYP and HF predict the BLA to be 1.13% and 3.29% higher than CCSD, respectively, while the values used in Torelli and Mitas's QMC study are 4.61% lower. These differences in geometries would also contribute to the energy differences observed—likely in very large amount with the nonoptimized structures employed in the QMC study. The behavior of rCAM-B3LYP, B3LYP, LDA, and HF for fractional numbers of electrons is exemplified in Fig. 4 for C_{14}. The maximum pointwise deviation from straight line behavior is 3.7 kcal/mol for rCAM-B3LYP, 5.7 kcal/mol for HF, 9.3 kcal/mol for B3LYP, and 12.4 kcal/mol for LDA. The rCAM-B3LYP deviation is similar to what was observed in an investigation of the energetics of Diels–Alder reactions.[19]

Another manifestation of the DE of DFAs is an incorrect treatment of the HOMO-LUMO gap. Figure 5 shows the gap for the lowest energy structures as a function of the ring size for various functionals. For this system we do not have high level results to compare with but we can understand the trend[18] and observe that the rCAM-B3LYP results improve upon other functionals. LDAs and GGAs greatly underestimate the gap as the convex curve for $E(N)$ underestimates the slope of E when approaching N from below and overestimates from above. With HF, the slope from above is too

TABLE I. Ring size at which the angle-alternating structure transitions to either the bond-alternating (BA), aromatic (AR), or bond- and angle-alternating (BA&A) form for various exchange-correlation functionals.

Functional	Transition
rCAM-B3LYP	$C_{10}\rightarrow C_{14}$ (BA)
CAM-B3LYP	$C_{10}\rightarrow C_{14}$ (B&AA)
B3LYP and PBE0	$C_{18}\rightarrow C_{22}$ (BA)
BLYP, PBE, and LDA[a]	$C_{10}\rightarrow C_{14}$ (AR)
QMC[b]	$C_{10}\rightarrow C_{14}$ (BA)

[a]A further transition to bond alternation should occur for very large rings owing to a Peierls mechanism.
[b]Reference 10.

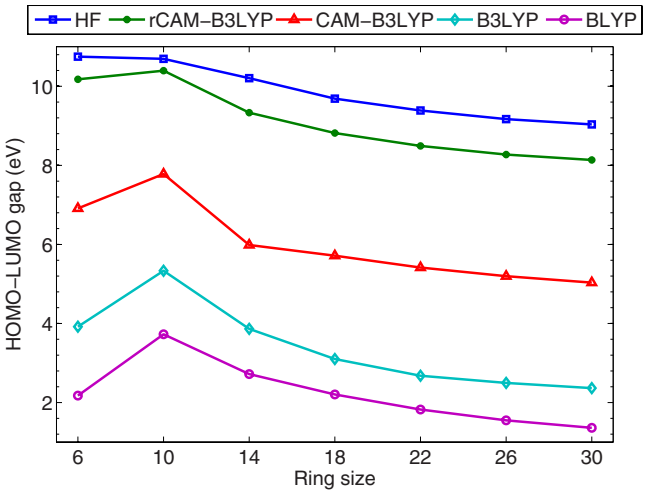

FIG. 4. HOMO-LUMO energy gap for the most stable structure of each functional as a function of the ring size for several exchange-correlation functionals. While not shown, the behavior of PBE0 is very similar to B3LYP; LDA (SVWN5) and PBE are very similar to BLYP.

small resulting in an overestimation of the gap. Hybrids improve upon GGAs as a result of their better $E(N)$ behavior from mixing with HF, with range-separated hybrids performing better than global hybrids.[22] From this argument we know that the correct value of the gap will fall between HF and CAM-B3LYP values, and this is what we see with rCAM-B3LYP. In Sec. III where we consider isomers of C_{20}, the HOMO-LUMO gaps we obtain with rCAM-B3LYP will be seen to be in very good agreement with coupled cluster calculations.

We consider one more system for which many conflicting DFT calculations exist–the ring, bowl, and cage isomers of C_{20}. Numerous HF calculations predicted the ring isomer to be the most stable and the cage the least. Numerous LDA calculations predicted the opposite ordering.[16] PBE (Ref. 28) and BLYP[29–31] GGAs agreed with the HF ordering. The B3LPY hybrid yielded either the bowl or ring,[30,32] depending on the basis set employed, as being most stable and cage the

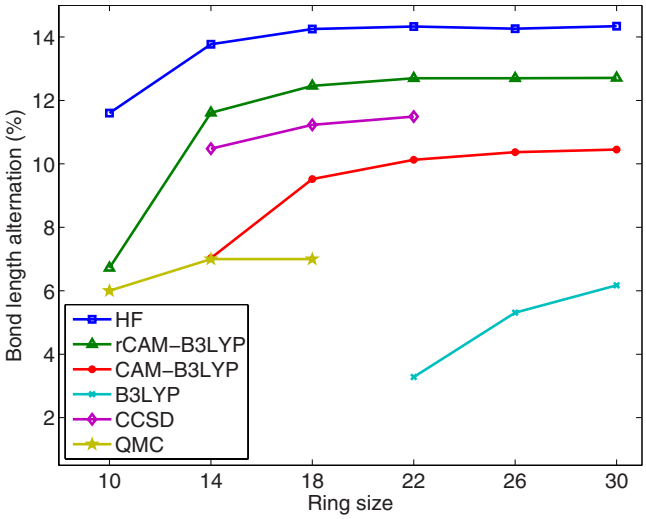

FIG. 5. BLA as a function of ring size. The CCSD results are from Arulmozhiraja and Ohno (Ref. 15) (employing the 6-31G* basis set) and QMC from Torelli and Mitas (Ref. 10).

TABLE II. Relative energies (eV) of the bowl, cage, and ring isomers of C_{20}.

Method	Bowl	Cage	Ring
CAM-B3LYP/cc-pVDZ	**0.0**	0.86	1.11
rCAM-B3LYP/cc-pVDZ	**0.0**	0.75	1.96
CCSD(T)/cc-pVDZ//PBE0/cc-pVTZ[a]	**0.0**	0.28	2.32
QMC[b]	**0.0**	2.15	1.06
B3LYP/cc-pVTZ[a]	0.0	1.89	**−0.90**

[a]Reference 32.
[b]Reference 30.

least, while B-PW91,[30] B3-PW91, and PBE1 (Ref. 32) favor the bowl structure. Local second order Moller–Plesset perturbation theory strongly favors the bowl structure[33] and is likely a reliable result given how well MP2 behaves with respect to the DE.[34]

The relative energies and the HOMO-LUMO gap of the isomers are shown in Tables II and III. Both rCAM-B3LYP and CAM-B3LYP predict the same energy ordering, with the bowl isomer being the most stable and the ring the least. This ordering is in agreement with the CCSD(T) calculations of An *et al.*,[32] where we also see reasonable agreement in the relative energies. Compared to their results, rCAM-B3LYP stabilizes the cage isomer by 0.38 eV and destabilizes the ring by 0.23 eV. The QMC results of Grossman *et al.*[30] also predict the bowl to be the most stable isomer; however, the ordering of the cage and ring is swapped. We are unable to pinpoint the reason for this discrepancy. Hybrid functionals do not predict a consistent ordering of these isomers[32] indicting that the DE is likely not the dominant source of error in determining the energy ordering. It is interesting to consider the HOMO-LUMO gap, for which the DE is of utmost importance. Here we see good agreement with CCSD(T) results: the rCAM-B3LYP results obtained directly from the HOMO and LUMO energies on the neutral species differ by 0.1 eV (bowl), 0.16 eV (cage), and 0.39 eV (ring) from CCSD(T)/cc-pVDZ//PBE0/cc-pVTZ calculations.[32] The coupled cluster calculations we compare with employ a double zeta basis set, owing to computational limitations. This is likely too small for the energies to be well converged, and some care in interpreting the extent of agreement must be made.

III. CONCLUDING REMARKS

In summary, we considered a structural manifestation of the DE of density functional approximations. The C_{4N+2} rings provide a transparent picture of such effects on molecular structures owing to the presence of a second order Jahn–Teller effect competing with Hückel aromaticity to produce a range of structures representing both localized and delocalized densities. Employing the rCAM-B3LYP exchange-correlation functional, which was designed to have minimal DE, we correctly predicted the geometric transition between angle alternation and bond alternation between C_{10} and C_{14}. rCAM-B3LYP along with MCY3 is the first functional designed with the intent to minimize the DE and the first to correctly describe this transition, emphasizing the importance of such functionals when studying systems where sta-

TABLE III. HOMO-LUMO gap (eV) of the bowl, cage, and ring isomers of C_{20}.

	$\varepsilon_{HOMO}-\varepsilon_{LUMO}$			$I-A$ [a]		
Method	Bowl	Cage	Ring	Bowl	Cage	Ring
CAM-B3LYP/cc-pVDZ	6.40	4.23	4.57	7.68	5.77	5.77
rCAM-B3LYP/cc-pVDZ	9.49	6.84	7.61	8.47	5.71	6.84
CCSD(T)/cc-pVDZ//PBE0/cc-pVTZ[b]				9.59	6.68	7.22
CCSD(T)/cc-pVDZ//MP2/cc-pVDZ[b]				9.13	6.38	6.52
B3LYP/cc-pVTZ[b]	3.72	1.93	2.07			

[a]Calculated as $[E_v(N-1)-E_v(N)]-[E_v(N)-E_v(N+1)]$.
[b]Reference 32.

bilization mechanisms compete with Hückel aromaticity. Our results overstabilize the structures with respect to QMC; however, good agreement with CCSD is achieved. We also applied the rCAM-B3LYP functional to the bowl, cage, and ring isomers of C_{20} where the DE is also relevant. We obtained an energy ordering in agreement with CCSD but not with QMC. The HOMO-LUMO gaps we obtained agree well with CCSD calculations owing to the improved straight line behavior of the total energy. Our results highlight the need of an unbiased DFT treatment of electron delocalization as a key for reliable prediction of relative stability and hence the structures of complex molecules where many structure stabilization mechanisms exist.

ACKNOWLEDGMENTS

Support from the National Science Foundation is greatly appreciated.

[1] P. Mori-Sánchez, A. J. Cohen, and W. Yang, Phys. Rev. Lett. **100**, 146401 (2008).
[2] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science **321**, 792 (2008).
[3] Y. Zhang and W. Yang, J. Chem. Phys. **109**, 2604 (1998).
[4] D. Jacquemin, A. Femenias, H. Chermette, I. Ciofini, C. Adamo, J.-M. André, and E. A. Perpète, J. Phys. Chem. A **110**, 5952 (2006).
[5] D. Jacquemin, E. A. Perpète, G. Scalmani, M. J. Frisch, and R. Kobayashi, J. Chem. Phys. **126**, 144105 (2007).
[6] M. D. Wodrich, C. Corminboeuf, and P. von Ragué Schleyer, Org. Lett. **8**, 3631 (2006).
[7] Y. Zhao, N. E. Schultz, and D. G. Truhlar, J. Chem. Theory Comput. **2**, 364 (2006).
[8] A. D. Becke and E. R. Johnson, J. Chem. Phys. **127**, 124108 (2007).
[9] A. V. Krukau, G. E. Scuseria, J. P. Perdew, and A. Savin, J. Chem. Phys. **129**, 124103 (2008).
[10] T. Torelli and L. Mitas, Phys. Rev. Lett. **85**, 1702 (2000).
[11] J. Hutter, H. P. Luethi, and F. Diederich, J. Am. Chem. Soc. **116**, 750 (1994).
[12] J. M. L. Martin, J. El-Yazal, and J. P. François, Chem. Phys. Lett. **242**, 570 (1995).
[13] M. Saito and Y. Okamoto, Phys. Rev. B **60**, 8939 (1999).
[14] S. Sen, P. Seal, and S. Chakrabarti, Phys. Rev. B **73**, 245401 (2006).
[15] S. Arulmozhiraja and T. Ohno, J. Chem. Phys. **128**, 114301 (2008).
[16] S. Sokolova, A. Lüchow, and J. B. Anderson, Chem. Phys. Lett. **323**, 229 (2000).
[17] P. Mori-Sánchez, Q. Wu, and W. Yang, J. Chem. Phys. **119**, 11001 (2003).
[18] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Phys. Rev. B **77**, 115123 (2008).
[19] E. R. Johnson, P. Mori-Sánchez, A. J. Cohen, and W. Yang, J. Chem. Phys. **129**, 204112 (2008).
[20] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Jr., Phys. Rev. Lett. **49**, 1691 (1982).
[21] W. Yang, Y. Zhang, and P. W. Ayers, Phys. Rev. Lett. **84**, 5172 (2000).
[22] A. J. Cohen, P. Mori-Sánchez, and W. Yang, J. Chem. Phys. **126**, 191109 (2007).
[23] E. J. Bylaska, J. H. Weare, and R. Kawai, Phys. Rev. B **58**, R7488 (1998).
[24] K. E. Yousaf and P. R. Taylor, Chem. Phys. **349**, 58 (2008).
[25] J. D. Watts and R. J. Bartlett, J. Chem. Theory Comput. **190**, 19 (1992).
[26] T. Yanai, D. P. Tew, and N. C. Handy, Chem. Phys. Lett. **393**, 51 (2004).
[27] DALTON, a molecular electronic structure program, Release 2.0, 2005, see http://www.kjemi.uio.no/software/dalton/dalton.html.
[28] G. Galli, F. Gygi, and J.-C. Golaz, Phys. Rev. B **57**, 1860 (1998).
[29] K. Raghavachari, D. L. Strout, G. K. Odom, G. E. Scuseria, J. A. Pople, B. G. Johnson, and P. M. W. Gill, Chem. Phys. Lett. **214**, 357 (1993).
[30] J. C. Grossman, L. Mitas, and K. Raghavachari, Phys. Rev. Lett. **75**, 3870 (1995).
[31] Z. Wang, P. Day, and R. Pachter, Chem. Phys. Lett. **248**, 121 (1996).
[32] W. An, Y. Gao, S. Bulusu, and X. C. Zeng, J. Chem. Phys. **122**, 204109 (2005).
[33] R. B. Murphy and R. A. Friesner, Chem. Phys. Lett. **288**, 403 (1998).
[34] A. J. Cohen, P. Mori-Sánchez, and W. Yang, J. Chem. Theory Comput. **5**, 786 (2009).