Theor Chem Acc (2011) 130:491–500
DOI 10.1007/s00214-011-1029-2

REGULAR ARTICLE

“Gold standard” coupled-cluster study of acetylene pentamers and hexamers via molecular tailoring approach

Anuja P. Rahalkar · Brijesh K. Mishra ·
V. Ramanathan · Shridhar R. Gadre

Received: 4 May 2011 / Accepted: 26 August 2011 / Published online: 15 September 2011

Abstract Due to high scaling order of MP2 and CCSD(T) methods, it is either difficult or at times even impossible to treat even moderately sized molecular systems with elaborate basis sets such as aug-cc-pVXZ ($X = D, T, Q$). In the present work, several structures of acetylene pentamers and hexamers are explored at MP2 and CCSD(T) levels of theory as prototypical examples of clusters bound by CH$\cdots\pi$ interactions. To enable this investigation, fragment-based method Molecular Tailoring Approach (MTA) is employed. It is shown that these acetylene assemblies can be treated with substantial reduction in computational resources and time, yet retaining a sub-millihartree accuracy in the energy. Further, using standard extrapolation methodologies, stabilization energies at the complete basis set limit of the acetylene clusters under consideration are determined at MP2 and CCSD(T) levels of theory. To test out the feasibility of treating a large cluster at MP2 level, a demonstrative calculation on a dodecamer of acetylene is reported.

Keywords Møller-Plesset second-order perturbation (MP2) theory · Coupled-cluster with single and double and perturbative triple excitations (CCSD(T)) · Acetylene · Molecular clusters · Complete basis set · Molecular tailoring approach (MTA)

1 Introduction

Weak interactions involving π-electron clouds of aromatic molecules with aliphatic/aromatic substrates have been attracting researchers from various fields in the past few years [1–3]. These weak interactions are found to play a significant role in various chemical systems and in dynamic processes. Among the weak interactions, C–H$\cdots\pi$ interaction is ubiquitous and has a major influence in both structural and functional aspects of biological macromolecules [4, 5], crystal packing [6], formation and stability of molecular clusters [7, 8], and self-assembly of supramolecules [9] and nanostructures [10]. In several systems, intramolecular C–H$\cdots\pi$ interaction is also a common feature [11]. These interactions are deemed to be distinct from the dispersion interactions, and they are placed between hydrogen bonds and dispersion interactions [12] purely from stability categorization viewpoint.

Exhaustive research work has been carried out on molecular clusters having $\pi\cdots\pi$ and C–H$\cdots\pi$ interactions in the last two decades [13–15]. Acetylene and benzene molecules are considered to be prototype for non-aromatic and aromatic π-systems, respectively. In both acetylene and benzene clusters, two types of interactions are present, i.e., $\pi\cdots\pi$ and C–H$\cdots\pi$. There is a striking similarity

---

Dedicated to Professor Shigeru Nagase on the occasion of his 65th birthday and published as part of the Nagase Festschrift Issue.

A. P. Rahalkar · S. R. Gadre (✉)
Department of Chemistry, Indian Institute of Technology,
Kanpur 208016, India
e-mail: gadre@iitk.ac.in

A. P. Rahalkar
Department of Chemistry, University of Pune,
Ganeshkhind Pune, Pune 411007, India

B. K. Mishra
Environmental Chemistry Modeling Laboratory,
GR C2 544, Swiss Federal Institute of Technology (EPFL),
Lausanne, Switzerland

V. Ramanathan
3rd Institute of Physics, University of Stuttgart,
57 Pfaffenwaldring, 70569 Stuttgart, Germany

between the two most stable structures of acetylene dimer and benzene dimer [13, 16]. Several studies are available in the literature for acetylene dimer, trimer, and tetramer complexes [16–19]. The two most stable structures for acetylene dimer are *T*-shaped (C_{2v} symmetry) and parallel-displaced (C_{2h} symmetry). The most stable structures for acetylene trimer and tetramer are the cyclic ones with C_{3h} and C_{4h} symmetry, respectively [16–19]. There are a limited number of studies available for acetylene pentamers and clusters higher than pentamer. Yu et al. reported two iso-energetic structures for acetylene pentamer, cyclic-planar (C_{5h} symmetry), and non-planar ring structure using the HF method [20]. Takeuchi [21] generated the geometries of acetylene clusters up to 55 acetylene molecules using intermolecular potential developed by Garrison and Sandler [22]. For the acetylene clusters (dimers to pentamers), a detailed theoretical investigation was recently carried out by Kim and co-workers employing high-level computational methods, such as the Møller-Plesset second-order perturbation (MP2) theory and coupled-cluster with single, double and perturbative triple excitations CCSD(*T*), extrapolated to complete basis set (CBS) limit [16].

Experimental measurement of stabilization energy for all the weakly bound systems is not readily possible, even with the state-of-the-art techniques. Quantum chemical calculations thus play a major guiding role in understanding the structures and stabilization energies of molecular aggregates. The problem of choosing a reliable level of theory for predicting the stabilization energy for weakly bound clusters often adds to the difficulty. Over the years, it has been shown and now well accepted that the Hartree–Fock (HF) and conventional density functional theory (DFT) based methods are inadequate for accurately quantifying the van der Waals interactions, in particular $\pi \cdots \pi$ and C–H$\cdots\pi$ ones [23]. Some of the recently developed DFT methods by Truhlar's and Grimme's group, for example, M06-2X and B2-PLYP, appear promising. However, they require further substantiation, especially when used for weakly bound molecular clusters [24, 25]. Consequently, MP2 theory, which incorporates correlation energy explicitly, has been used extensively for predicting reliable stabilization energy for such interactions. However, the major drawback with MP2 theory is that it overestimates the numerical value of the stabilization energy even with large enough basis sets such as aug-cc-p*VDZ* and aug-cc-p*VTZ* [14, 16]. Another problem associated with the MP2-based computation is the basis set superposition error (BSSE). The BSSE correction techniques, such as the counterpoise correction method, although commonplace, are nevertheless debatable, especially when it comes to weakly bound molecular clusters. Therefore, the MP2 prediction of stabilization energy for such weak interactions is rendered reliable only at the CBS limit; this becomes computationally prohibitive even for moderate-sized clusters due to high scaling order of the MP2 method, viz. $O(N^5)$. On the other hand, the application of highly accurate coupled-cluster methods such as CCSD(*T*), known as "gold standard" in quantum chemistry, which scale as $O(N^7)$, is still restricted to small molecules and clusters, owing to its high computational cost. In the recent years, the estimated CCSD(*T*)/CBS values are attaining the status of benchmark methods for the systems incorporating $\pi \cdots \pi$ and C–H$\cdots\pi$ contacts [14, 16]. Estimated CCSD(*T*)/CBS calculations rely on the assumption that the difference between the MP2/double-zeta and MP2/CBS energies would be very similar to that of the CCSD(*T*)/double-zeta and CCSD(*T*)/CBS ones. The extrapolated CCSD(*T*)/CBS value is obtained using a two-point extrapolation method, i.e., employing the CCSD(*T*)/aug-cc-p*VXZ* and CCSD(*T*)/aug-cc-p$V(X+1)Z$ values, where the cardinal number *X* is 2, 3, and 4 for *D*, *T*, and *Q*, respectively [26]. Nevertheless, it must also be noted here that the CCSD(*T*)/CBS method is not affordable for large clusters due to high computational cost.

The use of correlated ab initio methods is always restricted to small/medium sized molecular systems due to their formidable non-linear scaling. There are a few attempts in the literature to overcome this problem by divide-and-conquer (DC)-type strategies. The earliest attempt [27] in this direction was by Christoffersen et al. in 1972, which aimed at studying the strengths of hydrogen bonding in polyglycine and similar large systems. Later, Yang et al. proposed [28] a density matrix divide-and-conquer method (named as DC method) and applied it within the HF and DFT frameworks. Gadre and co-workers formulated [29, 30] the molecular tailoring approach (MTA) for evaluating one-electron properties of large molecules. This method was further extended to perform geometry optimization [31], calculate Hessian matrix and vibrational spectra [32] and molecular orbitals [33] of spatially extended large molecules. Another DC-type method viz. fragment molecular orbital (FMO) developed [34, 35] by Kitaura and co-workers has gained popularity during the last decade. This method scissors the molecule based on the concept of many-body interactions and is known to work exceptionally well for biomolecules. However, FMO has not been developed as a method applicable to general class of molecules. Also, if a covalent bond is cut within FMO, a basis set with diffuse functions cannot be employed. The DC method of Nakai et al. [36], systematic molecular fragmentation [37], molecular fractionation with conjugated caps [38], and generalized energy-based fragmentation [39] are some other similar DC-based methods reported in the literature. Except for the DC method by Nakai, all of these methods are again based on many-body analysis and not applicable to a general class of molecules. In contrast, being a more general method, MTA is applicable [40, 41] to any large molecular system without any restrictions on the level of theory and the basis set employed. This approach uses a distance criterion for fragmentation. A noteworthy study [42] on highly conjugated systems including a small model of graphene is a recent application of MTA.

Molecular Tailoring Approach (MTA) [29–33, 40–45] has been employed in the present work to perform calculations on acetylene clusters at MP2/CBS and CCSD(*T*)/CBS levels of theory. MTA has been shown to be a reliable and relatively less expensive computational protocol to handle large molecular systems and clusters at ab initio level using off-the-shelf hardware [31, 32, 42–45]. This method was earlier successfully applied to study the energetics and growth pattern of large boric acid nanotubes and benzene clusters [43–45]. Also, by performing standard calculations, it was demonstrated that the error introduced by MTA was only ~0.1 mH. Having validated the reliability of MTA for the benzene clusters, we proceed ahead to explore its applicability to study the acetylene clusters. Structurally, acetylene is an aliphatic analogue of benzene. Although acetylene is the smallest organic molecule where the $\pi\cdots\pi$ interaction and C–H$\cdots\pi$ interaction compete with each other in stabilizing the clusters, until recently, conclusive predictions were not available on the stabilization energy of larger acetylene clusters as mentioned above. Our recent success with the MTA method for benzene clusters [44] prompted us to investigate the pentamers and hexamers of acetylene. The choice of pentamer is made for re-validating the accuracy of MTA method, so that the accuracy of the results obtained for the hexamers can be considered unequivocal. To the best of our knowledge, high-level ab initio theoretical calculations for higher ($\geq$6 monomers) oligomers of acetylene are not available in the literature till date.

The main objective of the present work is to demonstrate the utility of MTA for handling large clusters at better correlated ab initio levels of MP2 and CCSD(*T*) and determining their stabilization energies at the CBS limit.

2 Methodology and computational details

In the present work, electronic structures of a few energetically favorable structures of acetylene (ac) pentamer and hexamer are investigated using quantum chemical methods. Initial geometries of these clusters are created using visualization software UNIVIS [46]. These structures are optimized at MP2/aug-cc-p*VDZ* level of theory using MTA employing Gaussian 09 [47] package at the back-end. The MTA-optimized structures of $(ac)_5$ are further subjected to geometry optimization using conventional method by Gaussian 09 package. Being simple structures from the fragmentation point of view, these are manually fragmented using visualization software MeTAStudio [48]. The main fragments of these are trimeric and tetrameric units depending upon the structure. A schematic illustration of fragmentation is given in Fig. 1. A prototypical $(ac)_6$ (Hex4) is broken into a set of 3 tetrameric units in order to capture the significant 2-, 3-, and 4-body contributions as per many-body energy decomposition analysis. A brief summary of MTA is given below, although the conceptual and algorithmic details of MTA may be found in Refs. [31, 32, 40, 41].

As seen in the previous section, MTA is a linear scaling method developed for handling large molecules and molecular clusters at ab initio level of theory. This method is based on a divide-and-conquer-type algorithm. The basic idea of MTA is to break a large system (parent molecule/cluster) into a set of smaller, manageable sub-problems (fragments), whose solutions are employed for estimating the result of the main system. Whenever a molecule is cut at covalent bonds, dummy hydrogen atoms are added to satisfy the unsaturated valencies. In case of molecular clusters, only intermolecular weak bonds are broken, and thus, no dummy atoms are added. For a system under consideration, many fragmentation schemes are possible, as the fragmentation in MTA is purely based on a distance criterion. The details of fragmentation procedure and criteria to choose an optimum fragmentation scheme are given in Ref. [31, 32, 41]. Ab initio calculations at the desired level of theory are performed on the fragments, and the energy or any property for the parent system is estimated using corresponding quantities of individual fragments. Cardinality-based expressions for energy and other properties such as gradients, Hessian matrix, and density matrix are generated employing set inclusion–exclusion

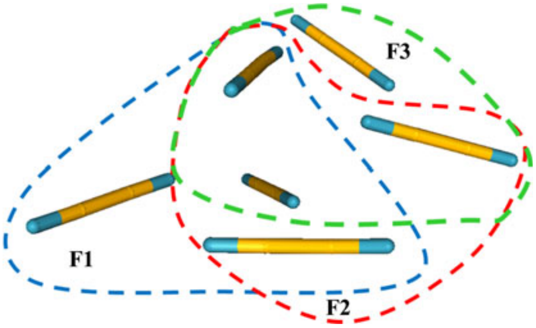

Fig. 1 Illustration of fragmentation of a prototypical acetylene hexamer $(ac)_6$ cluster into a set of tetrameric fragments viz. *F1*, *F2*, and *F3* denoted by *blue*, *red*, and *green colors*, respectively. See text and Ref. [44] for details

principle. The generalized form of MTA-based energy estimate is given as

$$E = \sum E^{f_i} - \sum E^{f_i \cap f_j} + \cdots + (-1)^{k-1} \sum E^{f_i \cap f_j \cap \cdots \cap f_k} \tag{1}$$

where E^{f_i} denotes energy of fragment f_i, $E^{f_i \cap f_i}$ denotes energy of the overlap of fragments fi and fj and so on. Similarly, generalized equation for gradient of energy with respect to a nuclear coordinate $X\mu$ can be written as

$$\frac{\partial E}{\partial X_\mu} = \sum \frac{\partial E^{f_i}}{\partial X_\mu^{f_i}} - \sum \frac{\partial E^{f_i \cap f_i}}{\partial X_\mu^{f_i \cap f_i}} + \cdots + (-1)^{k-1} \sum \frac{\partial E^{f_i \cap f_j \cap \cdots \cap f_k}}{\partial X_\mu^{f_i \cap f_j \cap \cdots \cap f_k}} \tag{2}$$

where $\frac{\partial E}{\partial X_\mu}$ and $\frac{\partial E^{f_i}}{\partial X_\mu^{f_i}}$ respectively denote the derivative of the MTA-energy and the energy of fragment f_i with respect to coordinates of atom μ and so on.

This method is now well established with benchmarks reported on various structurally diverse molecules and clusters for geometry optimization, Hessian and IR spectra calculation as well as evaluation of one-electron properties such as molecular electrostatic potential (MESP). Applicability of MTA at any level of theory employing any popularly available basis set makes it adaptable to newer faster ab initio programs as well as new methods (such as RI-MP2) or newer set of functionals (such as M05 and M06). In principle, any available ab initio suite of program can be run at the back-end driven by the MTA plugin utility. Currently, the MTA plugin utility is extended for GAMESS [49] and Gaussian [47] packages at the back-end. Calculations at MP2 level of theory, reported in the present work, are performed using Gaussian 09 package driven by MTA plugin utility. Further, all of the CCSD(*T*) computations were performed with CFOUR [50] package via MTA as CFOUR has been specially designed to perform efficient coupled-cluster calculations.

The final geometries of $(ac)_n$, $n = 5, 6$ at MP2/aug-cc-p*VDZ* (cf. Figs. 2, 3) are further subjected to MTA-based single-point energy calculations at MP2 and CCSD(*T*) method employing aug-cc-p*VDZ* and aug-cc-p*VTZ* basis sets. MTA-based single-point energy calculations are also performed at MP2/aug-cc-p*VQZ*. To illustrate that any higher cluster can be handled by MTA with a similar effort as that required for $(ac)_5$ and $(ac)_6$, a $(ac)_{12}$ structure chosen as an illustration, is optimized via MTA at MP2/aug-cc-p*VDZ* level of theory and subjected to a single-point energy calculation employing aug-cc-p*VTZ* basis set. The basis sets aug-cc-p*VXZ* ($X = D, T, Q$) are abbreviated henceforth in this article as XZ ($X = D, T, Q$). The error (ϵ in mH) in MTA-based energy estimate (E_{MTA}) is determined wherever possible, as follows:

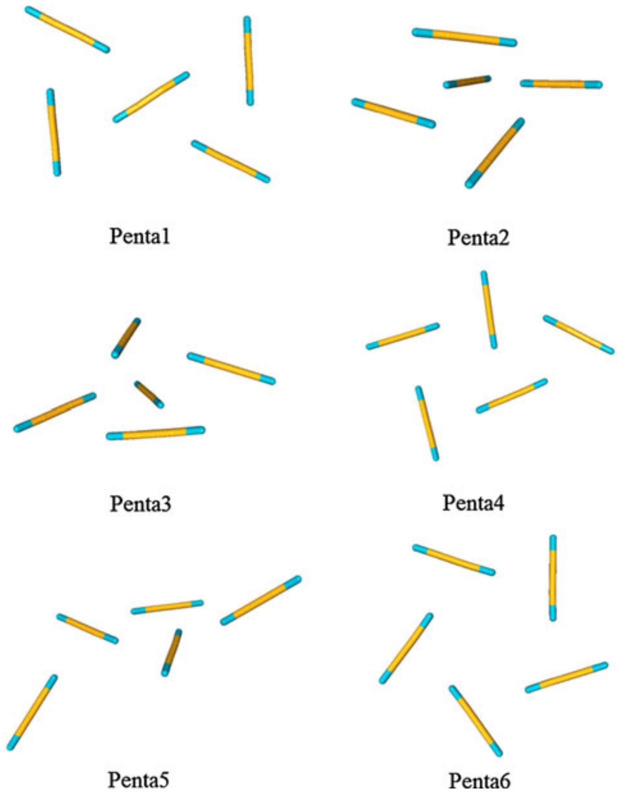

Fig. 2 Acetylene pentamer $(ac)_5$ structures, optimized at MP2/aug-cc-p*VDZ* level of theory. See Table 1 for the corresponding energies and text for details

$$\in = (E_{\mathrm{MTA}} - E_{\mathrm{Act}}) \times 1000 \tag{3}$$

where E_{Act} denotes actual energy of the cluster at the same level of theory. The stabilization energy (ΔE in kcal/mol) of cluster $(ac)_n$ is determined as

$$\Delta E = (E_{\mathrm{MTA}}/E_{\mathrm{Act}} - n \times E_{\mathrm{monomer}}) \times 627.51 \tag{4}$$

where E_{monomer} denotes the energy of acetylene monomer at the same level of theory and wherever possible, E_{Act} replaces E_{MTA} in Eq. 4.

In order to eliminate the basis set effect in terms of the basis set superposition error, the stabilization energy of a cluster is extrapolated to the complete basis set limit (CBS) at MP2 and CCSD(*T*) levels of theory using Helgaker's two-point formula [26]. Helgaker and co-workers have shown that the correlation contribution follows an X^{-3} form for aug-cc-p*VXZ* basis set. Thus, following formulae arise for CBS limit (CBS(*DT*) and CBS(*TQ*)) using *DZ*, *TZ* and *TZ*, *QZ* basis sets, respectively:

$$\mathrm{CBS}(DT) = [27 \times \Delta E(TZ) - 8 \times \Delta E(DZ)]/19 \tag{5}$$

$$\mathrm{CBS}(TQ) = [64 \times \Delta E(QZ) - 27 \times \Delta E(TZ)]/37 \tag{6}$$

where ΔE (XZ), $X = D, T, Q$ stands for stabilization energy of a cluster with XZ basis set. Thus, notations MP2/

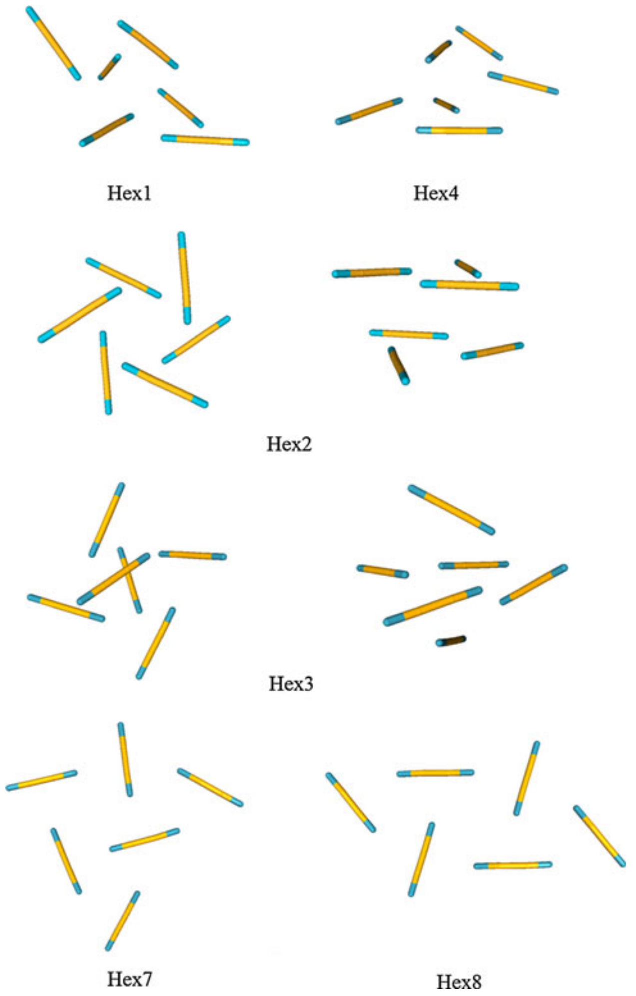

Fig. 3 A few most stable acetylene hexmer structures $(ac)_6$, MTA-optimized at MP2/aug-cc-p*VDZ* level of theory. Two views are provided for Hex2 and Hex3 for the sake of clarity. See text for further details and Table 3 for energies

CBS(*TQ*) and CCSD(*T*)/CBS(*DT*) used in forthcoming section denote stabilization energies of a cluster at MP2/CBS limit and CCSD(*T*)/CBS limit using the *XZ* basis sets indicated in the parentheses.

For large clusters, computations using the CCSD(*T*) method employing the *TZ* basis set could not be performed. Therefore, in such cases, stabilization energy at CCSD(*T*)/CBS is estimated as

$$\mathrm{CCSD}(T)/\mathrm{CBS(EST)} = \mathrm{MP2}/\mathrm{CBS}(TQ) + \left[\Delta E_{\mathrm{CCSD}(T)/DZ} - \Delta E_{\mathrm{MP2}/DZ}\right] \tag{7}$$

This is based on an assumption that the difference between the stabilization energies of a cluster at CCSD(*T*) and MP2 levels of theory is nearly constant for any basis set employed.

3 Results and discussion

Six different clusters of $(ac)_5$ are optimized employing MTA, followed by the actual run. It is worth mentioning that both planar and non-planar clusters are considered in the present study, in contrast to the other previous studies [16, 19] wherein the emphasis is laid only on the planar clusters. The total electronic energies of these six $(ac)_5$ obtained using MTA are listed in Table 1. Corresponding actual calculations are performed for benchmarking purpose. It must also be noted that the energies are reported in the atomic units (Hartree) and the errors (ϵ) are reported in mH units (cf. Eq. 3), which in fact very appropriately emphasizes the accuracy of MTA method. Average absolute error ($\bar{\epsilon}$) for the MP2/*DZ*, MP2/*TZ*, MP2/*QZ*, and CCSD(*T*)/*DZ* level turns out to be 0.40, 0.36, 0.37, and 0.25 mH, respectively, bringing out excellent agreement of MTA results with those obtained from the standard methods. The largest error produced by MTA method is 0.66 mH, which is still less than 0.5 kcal/mol. In the recent study, on the acetylene clusters [16], the highest level of theory used is the CCSD(*T*) with *DZ* basis set, which is understandable since enormous computational resources are required at the CCSD(*T*)/*TZ* level for acetylene pentamer. However, with the powerful tool of MTA at hand, it is indeed possible to carry out these calculations by breaking the pentamers into set of trimer and/or tetramer fragments, which are manageable at CCSD(*T*)/*TZ* level of theory. The error column for the CCSD(*T*)/*TZ* is missing from Table 1 as actual calculations at this level could not be performed on the available hardware. Overall, the error values listed in Table 1 do demonstrate high accuracy of the results.

Table 2 reports the stabilization energies for the different pentamer clusters under consideration. Recently, Kim and his co-workers reported [16] the two most stable pentamers of C_1 and C_s symmetry (5C1 and 5Cs, respectively as per the nomenclature in Ref. [16]). As mentioned above, in the present study, the stabilization energies are computed using two additional sets of calculations (MP2/*QZ* and CCSD(*T*)/*TZ*), which were not carried out previously [16]. Interestingly, both the benchmarks, that is, the CCSD(*T*)/CBS(EST) and CCSD(*T*)/CBS(*DT*), have values very close to each other. Therefore, any of the two can be considered as the best benchmark. In the following discussion, we have accepted the CCSD(*T*)/CBS(EST) as the best value. We would like to emphasize that our focus is mainly on the non-planar clusters against the planar clusters studied in Ref. [16]. Two of the planar $(ac)_5$ structures, viz. Penta1 and Penta6, are identical to those reported in Ref. [16]. From Table 1, it may be seen that the stabilization energy values for these two clusters are in good agreement with the ones reported in the above reference. It is to be noted that the earlier reported values are subjected

Table 1 MTA-based energies (E_{MTA}) in hartree for acetylene pentamers, $(\mathrm{ac})_5$, at MP2 level employing aug-cc-pVXZ ($X = D$, T, Q) and CCSD(T) level using aug-cc-pVXZ ($X = D$ and T) basis sets

System	MP2/*DZ*		MP2/*TZ*		MP2/*QZ*		CCSD(*T*)/*DZ*		CCSD(*T*)/*TZ*	
	E_{MTA}	$\bar{\epsilon}$	E_{MTA}	$\bar{\epsilon}$	E_{MTA}	$\bar{\epsilon}$	E_{MTA}	$\bar{\epsilon}$	E_{MTA}	$\bar{\epsilon}$
Penta1	−385.48786	0.47	−385.83492	0.29	−385.94065	0.24	−385.62821	0.26	−385.97317	–
Penta2	−385.48991	0.00	−385.83677	−0.01	−385.94243	−0.39	−385.62942	−0.06	−385.97409	–
Penta3	−385.48956	0.45	−385.83645	0.25	−385.94219	0.15	−385.62868	0.19	−385.97350	–
Penta4	−385.48944	−0.23	−385.83627	−0.31	−385.94176	−0.34	−385.62971	−0.18	−385.97441	–
Penta5	−385.48506	−0.59	−385.83251	−0.64	−385.93832	−0.46	−385.62567	−0.39	−385.97110	–
Penta6	−385.48587	0.64	−385.83303	0.64	−385.93866	0.66	−385.62648	0.39	−385.97159	–
$\bar{\epsilon}$	–	0.40	–	0.36	–	0.37	–	0.25	–	–

The errors (ϵ) in E_{MTA} are given in millihartree (mH), (cf. Eq. 3), $\bar{\epsilon}$ being the average of absolute errors. See text for details

Table 2 MTA-based stabilization energies ΔE of acetylene pentamers $(\mathrm{ac})_5$, for MP2/aug-cc-pVXZ ($X = D$, T, Q) and CCSD(T)/aug-cc-pVXZ ($X = D$, T) levels of theory

System	ΔE/MP2-aug-cc-pVXZ				ΔE/CCSD(T)-aug-cc-pVXZ			
	D	*T*	*Q*	CBS(*TQ*)	*D*	*T*	CBS(*DT*)	CBS(MP2-D)
Penta1	−14.36	−11.99	−11.03	−10.34	−13.43	−10.69	−9.54	−9.41
Penta2	−15.64	−13.15	−12.15	−11.42	−14.19	−11.27	−10.04	−9.97
Penta3	−15.42	−12.95	−12.00	−11.31	−13.72	−10.90	−9.71	−9.62
Penta4	−15.35	−12.83	−11.73	−10.92	−14.37	−11.47	−10.25	−9.95
Penta5	−12.60	−10.48	−9.57	−8.91	−11.83	−9.39	−8.36	−8.15
Penta6	−13.11	−10.80	−9.79	−9.04	−12.34	−9.70	−8.59	−8.28

CBS(*TQ*) and CBS(*DT*) denote corresponding stabilization energy at complete basis set limit employing 2-point extrapolation formula, while CBS(MP2-D) denotes that estimated by grafting CCSD(T) correction on MP2/CBS stabilization energy. All values are in kcal/mol. See text for details

to BSSE correction, whereas our values reported in Table 1 are not corrected for BSSE. It is well-known that as the limiting case of infinite basis set is approached, the BSSE tends to zero. This can be precisely observed in the current study of $(\mathrm{ac})_5$, where larger basis sets are used to extrapolate the stabilization energies to those at the basis set limit. Thus, it is safely assumed that while applying these CCSD(T)/CBS schemes, the correction for BSSE is no longer required. As seen from Table 2, Penta2 and Penta4 are the most stable clusters at the CCSD(T)/CBS(EST) limit, having stabilization energy values −9.97 and −9.95 kcal/mol, respectively. Penta2 is a non-planar structure (C_1 symmetry) exhibiting six C–H$\cdots\pi$ interactions as shown in Fig. 4. On the other hand, Penta4 is nearly a planar structure (C_1 symmetry) exhibiting again six C–H$\cdots\pi$ interactions. The next two most stable structures at the CCSD(T)/CBS(EST) limit are Penta3 (non-planar) and Penta1 (planar), respectively. These CCSD(T)/CBS limit stabilization energies for Penta2 and Penta4 clearly indicate that for the $(\mathrm{ac})_5$, both planar and non-planar structures are endowed with compatible stability. The clusters Penta5 and Penta6 are found to be the least stable cluster among those studied.

For $(\mathrm{ac})_6$, 8 different structures are scanned, out of which Hex1 and Hex2 are newly built, while Hex3 to Hex8 are generated by adding an acetylene monomer at intuitively suitable positions in the above mentioned $(\mathrm{ac})_5$ structures. Among the eight hexamer clusters, Hex6, Hex7, and Hex8 are generated from Penta4 resulting in a planar structure; the other five are non-planar. The accuracy of the fragmentation schemes is tested for MP2/aug-cc-pVXZ, $X = D$, and T levels where actual calculations are also performed for $(\mathrm{ac})_6$. As seen from Table 3, the maximum error involved in MTA calculations is 1 mH($\sim$0.6 kcal/mol) for the case of Hex5 at MP2/aug-cc-pVTZ level. Overall, the errors (ϵ, as per Eq. 3) in MTA-estimated energies are about 0.5 mH ($\sim$0.3 kcal/mol). The higher magnitude of ϵ in case of Hex5 can be attributed to its fragmentation into trimeric fragments. For the other $(\mathrm{ac})_6$ structures, trimeric as well as tetrameric fragments are employed whenever necessary.

Table 4 reports the MTA-based stabilization energies for all the $(\mathrm{ac})_6$ at MP2 as well as CCSD(T) levels of theory. Hex2 stands out as the most stable structure at MP2 and CCSD(T) levels of theories for all the basis sets examined with a stabilization energy of −13.06 kcal/mol

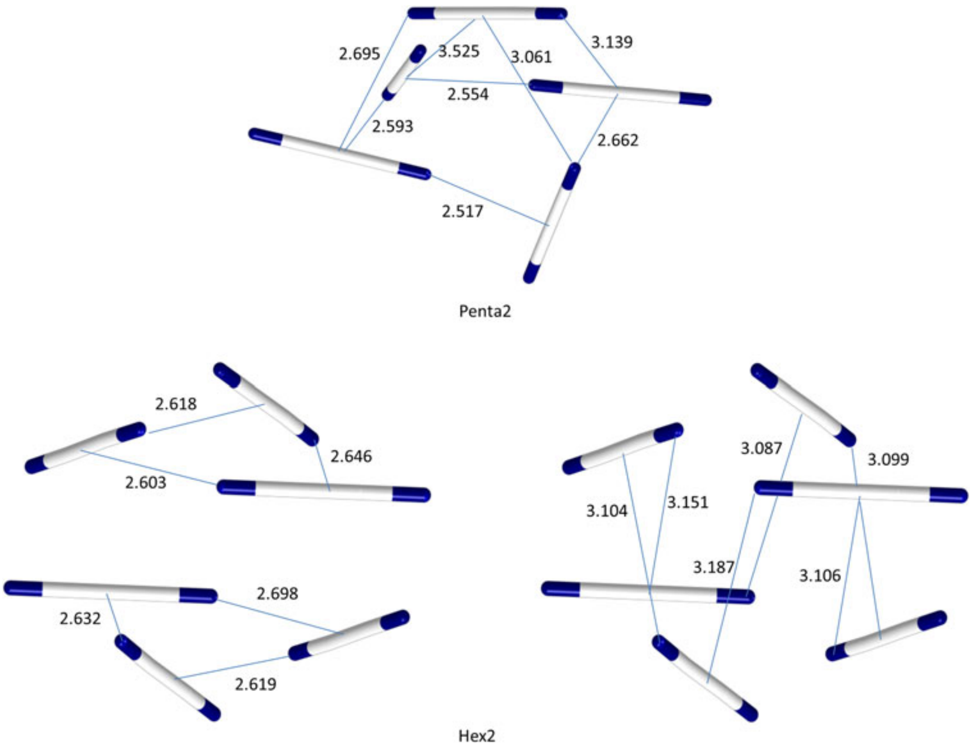

Fig. 4 Distances (in Å) for significant CH···π interactions in most stable pentamer (Penta2) and hexamer (Hex2) of acetylene. Two views are provided for Hex2 to show intra-planar and inter-planar interactions among acetylene units. See text for details

Table 3 MTA-based energies (E_{MTA}) for acetylene hexamers at MP2 level employing aug-cc-p*VXZ* ($X = D, T, Q$) and CCSD(*T*) level using aug-cc- p*VXZ* ($X = D$ and *T*) basis sets are reported in a.u.

System	MP2/*DZ*		MP2/*TZ*		MP2/*QZ*		CCSD(*T*)/*DZ*		CCSD(*T*)/*TZ*	
	E_{MTA}	$\bar{\epsilon}$	E_{MTA}	$\bar{\epsilon}$	E_{MTA}	$\bar{\epsilon}$	E_{MTA}	$\bar{\epsilon}$	E_{MTA}	$\bar{\epsilon}$
Hex1	−462.59068	0.33	−463.00649	−0.02	−463.13307	–	−462.75754	–	−463.17060	–
Hex2	−462.59231	0.79	−463.00747	0.64	−463.13391	–	−462.75873	–	−463.17096	–
Hex3	−462.59063	0.61	−463.00608	0.46	−463.13254	–	−462.75695	–	−463.16952	–
Hex4	−462.59066	0.51	−463.00621	0.24	−463.13278	–	−462.75705	–	−463.16952	–
Hex5	−462.58823	0.84	−463.00399	1.06	−463.13110	–	−462.75522	–	–	–
Hex6	−462.58827	0.47	−463.00431	0.22	−463.13090	–	−462.75641	–	−463.16988	–
Hex7	−462.58948	−0.45	−463.00525	−0.50	−463.13179	–	−462.75751	–	–	–
Hex8	−462.58928	−0.50	−463.00513	−0.70	−463.13172	–	−462.75733	–	–	–
$\bar{\epsilon}$	–	0.56	–	0.48	–	–	–	–	–	–

The errors (ϵ) in E_{MTA} are given in millihartree (mH) (cf. Eq. 3), $\bar{\epsilon}$ being the average of absolute errors. See text for details

at CCSD(*T*)/CBS(EST). This structure comprises of two acetylene trimer units (in most stable arrangement) stacked in staggered form (Fig. 3). As depicted in Fig. 4, 12 CH···π interactions are seen in Hex2. The stabilization energy for Hex2 at MP2/CBS limit is −15.42 kcal/mol, while that at CCSD(*T*)/CBS limit is −13.06 kcal/mol. The next three energetically favorable structures are Hex1, Hex7, and Hex8 at the CCSD(*T*)/CBS(EST), which are very close to each other in terms of stabilization energies. Hex5 is found to be the least stable among all.

In order to bring out a further example showing the utility of MTA in treating large molecular assemblies at MP2 level of theory, a cluster of 12 acetylene units (cf. Fig. 5) is chosen. The stabilization energies of the final MTA-optimized geometries at MP2/*DZ* and MP2/*TZ* are −51.64 and −42.72 kcal/mol, respectively. Extrapolating

Table 4 MTA-based stabilization energies ΔE of acetylene hexamers $(ac)_6$, for MP2/aug-cc-p*VXZ* ($X = D, T, Q$) and CCSD(T)/aug-cc-p*VXZ* ($X = D, T$) levels of theory

System	ΔE/MP2-aug-cc-p*VXZ*				ΔE/CCSD(T)-aug-cc-p*VXZ*			
	D	*T*	*Q*	CBS(*TQ*)	*D*	*T*	CBS(*DT*)	CBS(MP2-D)
Hex1	−20.52	−17.26	−15.93	−14.96	−18.43	−14.59	−12.97	−12.87
Hex2	−21.54	−17.88	−16.46	−15.42	−19.18	−14.82	−12.98	−13.06
Hex3	−20.49	−17.01	−15.60	−14.57	−18.06	−13.91	−12.16	−12.14
Hex4	−20.50	−17.09	−15.75	−14.78	−18.12	−13.91	−12.14	−12.40
Hex5	−18.98	−15.96	−14.70	−13.78	−16.97	–	–	−11.77
Hex6	−19.01	−15.89	−14.57	−13.60	−17.72	−14.14	−12.63	−12.32
Hex7	−19.77	−16.48	−15.13	−14.14	−18.41	–	–	−12.78
Hex8	−19.64	−16.41	−15.09	−14.12	−18.30	–	–	−12.78

CBS(*TQ*) and CBS(*DT*) denote corresponding stabilization energy at complete basis set limit employing 2-point extrapolation formula, while CBS(MP2-D) denotes that estimated by grafting CCSD(T) correction on MP2/CBS stabilization energy. All values are in kcal/mol. See text for details

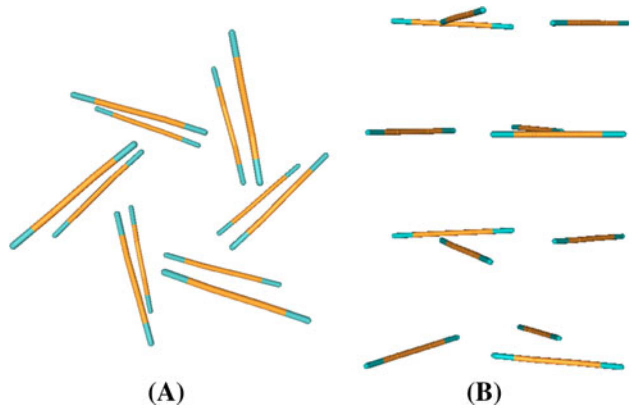

Fig. 5 Top view (**a**) and side view (**b**) of an illustrative case of MTA-optimized structure of acetylene dodecamer $(ac)_{12}$ at MP2/aug-cc-p*VDZ* level of theory. See text for details

these to CBS limit by employing Eq. 5, the stabilization energy at MP2/CBS(*DT*) turns out to be −38.96 kcal/mol. This example of $(ac)_{12}$ shows the applicability of the present methodology for treating any higher assemblies in similar economical way by breaking them into a set of trimer and/or tetramer fragments.

Apart from the feasibility of calculations on larger clusters, even with limited hardware, it is also important to assess the time advantage achieved due to MTA when the conventional calculation is possible. For instance, MTA-based calculation on Penta1 at CCSD(T)/*DZ* level of theory took ~10 h, while the actual calculation took ~100 h on identical hardware. Thus, a clear advantage factor of 10 is achieved by MTA. For the case of Penta4, MTA and actual timings are ~23 and ~100 h, with the advantage factor turning out to be ~4. For the any $(ac)_6$ structure, it is obvious that the T_r will increase, as the computational time for actual calculation will increase tremendously due to scaling of $O(N^7)$ of CCSD(T) method. But MTA time is expected to remain similar to that for $(ac)_5$, as all $(ac)_5$ and $(ac)_6$ are broken as trimers and/or tetramers!

For acetylene dimer, the most stable structures predicted in the literature are *T*-shaped and stacked-displaced. In higher acetylene clusters such as pentamers and hexamers, acetylene molecules tend to re-arrange themselves in three-dimensional space to maximize number of C–H···π contacts. Therefore, most of the interactions present are of type C–H···π interactions, though their orientation no longer remains as perfect *T*-shaped or stacked-displaced as in the case of the dimer. From the magnitude of the interaction energy listed for the acetylene pentamers and hexamers, it can be deduced that the bonding strength of C–H···π interactions lies between that of typical hydrogen bond (~5 kcal/mol) and purely dispersion-bound complexes of similar size (less than 1 kcal/mol).

Interestingly, the stabilization energies of several planar and non-planar configurations fall close to each other; hence, it is difficult to select the “global minimum” configuration or the most favorable ones among them. This study unfolds the possibility that as the number of acetylene molecules increase in the cluster, the number of potential configurations also increases.

It is a well-known fact that MP2 overestimates the stabilization energy. The present study also confirms that the MP2 method overestimates the interaction energy even when correlation consistent basis sets as large as aug-cc-p*VTZ* and aug-cc-p*VQZ* are employed. Hence, coupled-cluster calculations are essential to predict accurate stabilization energies for such systems. The present study has revealed that such studies are indeed feasible.

4 Concluding remarks

With the advent in computing hardware and software, it is now possible to handle large molecular system or clusters with computationally expensive but accurate methods such as MP2 and CCSD(T). However, due to non-linear scaling of these methods, there is a natural limit to the size of the clusters that can be treated with "black-box" packages of programs. Thus, a need is felt for linear scaling parallelization strategies or use of fragment-based methods such as MTA to enable the treatment of higher clusters on minimal hardware that is currently available to the user.

The earlier attempts [43–45] to treat molecular clusters via MTA at HF, DFT and MP2 levels of theory have been successful in achieving high accuracy at low computational cost. On the similar lines, the present work exploits the use of MTA for acetylene clusters at MP2 level of theory with correlation consistent basis sets and complete basis set limit. Special feature of the present study is extending out the applicability of MTA to CCSD(T) level of theory, which is *de facto* benchmark for studying the stabilization energies. There are reports in the literature [16] wherein (ac)$_5$ are treated at MP2 and CCSD(T)/DZ level of theory using the conventional methods. However, it would be rather difficult to treat (ac)$_5$ at CCSD(T)/TZ level of theory, and further, the computational requirements for hexamers at the similar level would make it practically impossible. The current investigation shows that on a cluster of off-the shelf computers, in principle, any higher cluster of acetylene can be treated. In general, 1.5 mH ($\sim$1 kcal/mol) can be taken as well-accepted chemical accuracy. But as the sophistication of level of theory grows, a need is felt to reduce the errors within 0.5 mH or even less. Thus, there is a further scope of improvement for the MTA-based investigation of clusters at MP2 and CCSD(T) level in improving the accuracy of the method.

In summary, the present study has clearly demonstrated that the MTA offers a solution for the treatment of large molecular clusters of chemical interest at the current "gold standard" of highly accurate methods such as CCSD(T).

Acknowledgments APR is thankful to the Council of Scientific and Industrial Research (CSIR), New Delhi, for award of research fellowship and Center for Development of Advanced Computing (C-DAC), Pune, for computational resources. Department of Science and Technology (DST), New Delhi, is gratefully acknowledged by SRG for the award of J. C. Bose fellowship.

References

1. Chenoweth K, Dykstra CD (2003) Theor Chem Acc 110:100
2. Kim KS, Tarakeshwar P, Lee JY (2000) Chem Rev 100:4145
3. Chourasia M, Sastry GM, Sastry GN (2011) Int J Biol Macromol 48:540
4. Burley SK, Petsko GA (1985) Science 229:23
5. Nishio M, Hirota M, Umezawa Y (1998) The CH/π interaction: evidence, nature, and consequences. Wiley, New York
6. Jones PG, Dix I, Hopf H (2007) Acta Cryst C 63:468
7. Řezáč J, Fanfrlík J, Salahub D, Hobza P (2009) J Chem Theory Comput 5:1749
8. Vaupel S, Brutschy B, Tarakeshwar P, Kim KS (2006) J Am Chem Soc 128:5416
9. Singh NJ, Lee HM, Hwang I-C, Kim KS (2007) Supramol Chem 19:321
10. Lee JY, Hong BH, Kim WY, Min SK, Kim Y, Jouravlev MV, Bose R, Kim KS, Hwang I-C, Kaufman LJ, Wong CW, Kim P, Kim KS (2009) Nature 460:498
11. Karaminkov R, Chervenkov S, Neusser HJ, Ramanathan V, Chakraborty T (2009) J Chem Phys 130:034301
12. Kenta S, Fujii A, Mikami N, Tsuzuki S (2007) J Phys Chem A 111:753
13. Sinnokrot OM, Sherrill DC (2006) J Phys Chem A 110:10656
14. Mishra BK, Samuel JA, Sathyamurthy N (2010) J Phys Chem A 114:9606
15. Sundararajan K, Sankaran K, Vishwanathan KS, Kulkarni AD, Gadre SR (2002) J Phys Chem A 106:1504
16. Karthikeyan S, Lee HM, Kim KS (2010) J Chem Theory Comput 6:3190
17. Shuler K, Dykstra CE (2000) J Phys Chem A 104:11522
18. Ohshima Y, Matsumoto Y, Takami M, Kuchitsu K (1988) Chem Phys Lett 147:1
19. Bone RGA, Amos RD, Handy NC (1990) J Chem Soc Faraday Trans 86:1931
20. Yu J, Shujun S, Bloor JE (1990) J Phys Chem 94:5589
21. Takeuchi H (2010) J Comp Chem 31:1699
22. Garrison SL, Sandler SI (2004) J Phys Chem B 108:18972
23. Mishra BK, Sathyamurthy N (2005) J Phys Chem A 109:6
24. Zhao Y, Truhlar DG (2008) Acc Chem Res 41:157
25. Goerigk L, Grimme S (2010) J Chem Theory Comput 6:107
26. Halkier A, Klopper W, Helgaker T, Jørgensen P, Taylor PR (1999) J Chem Phys 111:9157
27. Shipman LL, Christoffersen RE (1972) Proc Nat Acad Sci USA 69:3301
28. Yang W (1991) Phys Rev Lett 66:1432
29. Gadre SR, Shirsat RN, Limaye AC (1994) J Phys Chem 98:9165
30. Babu K, Gadre SR (2003) J Comput Chem 24:484
31. Ganesh V, Dongare RK, Balanarayan P, Gadre SR (2006) J Chem Phys 125:104109
32. Rahalkar AP, Ganesh V, Gadre SR (2008) J Chem Phys 129:234101
33. Rahalkar AP, Gadre SR (2011) J Chem Sci (in Press)
34. Kitaura K, Ikeo E, Asada T, Nakano T, Ubeyasi M (1999) Chem Phys Lett 313:701
35. Fedorov DG, Jensen JH, Deka RC, Kitaura K (2008) J Phys Chem A 111:6904
36. Kobayashi M, Akama T, Nakai H (2006) J Chem Phys 125:204106
37. Deev V, Collins MA (2005) J Chem Phys 122:154102
38. Zhang DW, Zhang JZH (2003) J Chem Phys 119:3599
39. Li W, Li S (2004) J Chem Phys 121:6649
40. Rahalkar AP, Katouda M, Gadre SR, Nagase S (2010) J Comput Chem 31:2405
41. Rahalkar AP, Yeole SD, Ganesh V, Gadre SR (2011) In: Zaleśny R, Papadopoulos MG, Mezey PG, Leszczynski J (eds) Linear scaling techniques in computational chemistry and physics. Springer, New York, p 199
42. Yeole SD, Gadre SR (2010) J Chem Phys 132:094102
43. Elango M, Subramanian V, Rahalkar AP, Gadre SR, Sathyaurthy N (2008) J Phys Chem A 112:7699
44. Mahadevi AS, Rahalkar AP, Gadre SR, Sastry GN (2010) J Chem Phys 133:164308

45. Gadre SR, Jovan JKV, Rahalkar AP (2010) J Chem Sci 122:1
46. Limaye AC, Gadre SR (2001) Curr Sci (India) 80:1296
47. The GAUSSIAN 09 package: Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA, Peralta Jr JE,Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, FarkasÖ, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, Revision A.1 Gaussian, Inc., Wallingford
48. The package MeTA Studio available at: http://code.google.com/p/metastudio/ (See: Ganesh V (2009) J Comp Chem 30:661)
49. The GAMESS package: Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su SJ, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347 For further details: http://www.msg.ameslab.gov/GAMESS/GAMESS.html
50. The CFOUR package: Auer AA, Bartlett RJ, Benedikt U, Berger C, Bernholdt DE, Bomble YJ, Cheng L, Christiansen O, Heckert M, Heun O, Huber C, Jagau T-C, Jonsson D, Jusélius J, Klein K, Lauderdale WJ, Matthews DA, Metzroth T, O'Neill DP, Price DR, Prochnow E, Ruud K, Schiffmann F, Schwalbach W, Stopkowicz S, Tajti A, Vázquez J, Wang F, Watts JD and the integral packages MOLECULE (Almlöf J and Taylor PR), PROPS (Taylor PR), ABACUS (Helgaker T, Jensen HJAa, Jørgensen P, and Olsen J), and ECP routines by Mitin AV and van Wüllen C