

## Molecular tailoring approach for exploring structures, energetics and properties of clusters

SHRIDHAR R GADRE\*, K V JOVAN JOSE and ANUJA P RAHALKAR

Department of Chemistry, University of Pune, Ganeshkhind, Pune 411 007

e-mail: gadre@chem.unipune.ernet.in

**Abstract.** Molecular Tailoring Approach (MTA) is a method developed for enabling *ab initio* calculations on prohibitively large molecules or atomic/molecular clusters. A brief review of MTA, a linear scaling technique based on set inclusion and exclusion principle, is provided. The Molecular Electrostatic Potential (MESP) of smaller clusters is exploited for building initial geometries for the larger ones, followed by MTA geometry optimization. The applications of MTA are illustrated with a few test cases such as  $(\text{CO}_2)_n$  and  $\text{Li}_n$  clusters employing Density Functional theory (DFT) and a nanocluster of orthoboric acid at the Hartree–Fock (HF) level. Further, a discussion on the geometries and energetics of benzene tetramers and pentamers, treated at the Møller–Plesset second order (MP2) perturbation theory, is given. MTA model is employed for evaluating some cluster properties viz. adiabatic ionization potential, MESP, polarizability, Hessian matrix and infrared frequencies. These property evaluations are carried out on a series of test cases and are seen to offer quite good agreement with those computed by an actual calculation. These case studies highlight the advantages of MTA model calculations vis-à-vis the actual ones with reference to the CPU-time, memory requirements and accuracy.

**Keywords.** Molecular clusters; linear scaling methods; molecular tailoring approach (MTA); Hartree–Fock (HF) method; density functional theory (DFT); Møller–Plesset second order (MP2) method; molecular electron density (MED); molecular electrostatic potential.

### 1. Introduction

A vast amount of research is being carried out for the design and synthesis of nanomaterials, which are suitable for a variety of technological applications. These nanomaterials are aggregates of atoms or molecules of nanometer scale and show altogether different properties from those of the isolated state and of the bulk. This can be attributed to the high surface area to volume ratio of the cluster as compared to that of the bulk as well as vastly different electronic properties.<sup>1</sup> It is possible to vary these peculiar electronic properties of nanomaterials by controlling their sizes. Thus, these materials are well-suited for applications in mechanical devices, electronics, optics, catalysts, etc.<sup>2</sup>

The development of reliable methods for modeling the energetics of the nanoclusters is a rapidly developing field. The starting point for understanding of cluster properties is study of its structure. The semi-empirical methods are extremely fruitful for selecting good candidates for subsequent *ab initio*

structural optimization.<sup>1</sup> It is indeed expensive to handle these clusters comprising of hundreds of atoms by the *ab initio* methods. The main hurdle of *ab initio* methods lies in their dependence on the number of basis functions ( $N$ ). It is known that the Hartree–Fock (HF) theory scales typically as  $O(N^3)$  for large systems, while the Møller–Plesset second order perturbation theory (MP2) goes as  $O(N^5)$ . Methods treating the electronic correlation in a better way, such as the coupled cluster (CC) theory, scale as even higher powers of  $N$ . Such a formidable scaling behaviour restricts the use of *ab initio* methods in optimizing the geometries of molecular systems typically a few hundreds of atoms at the HF level and to a few tens of atoms even for intermediate-quality correlated methods (e.g. MP2) with the contemporary off-the shelf hardware.

Many research groups across the globe have been actively working to find alternate methods to circumvent the above mentioned formidable scaling problem of the *ab initio* methods. The earliest attempt in this direction was reported by Christoffersen *et al.*<sup>3</sup> in 1972 for understanding the strength of hydrogen bonding in polyglycine. They also car-

\*For correspondence

ried out many studies on biological systems using their methodology. Later, Yang *et al*<sup>4,5</sup> proposed a density matrix divide-and-conquer (DC) method for electronic structure calculation of large molecules. The method is based on partitioning the density matrix and is applicable to both HF and DFT level of theories.

Subsequently, Gadre *et al*<sup>6</sup> formulated the Molecular Tailoring Approach (MTA) for evaluating one-electron properties of large molecular systems. The method was then developed further<sup>7</sup> for geometry optimization of large, spatially extended molecular systems. A similar DC-type method, viz. Fragment Molecular Orbital (FMO)<sup>8-15</sup> has been extensively developed by Kitaura's group, which has been incorporated in GAMESS<sup>16</sup> package. Nakai and co-workers have recently proposed a DC-based MP2 method (DCMP2).<sup>17-19</sup> In this method, correlation energy of the parent system is calculated using contributions from sub-system orbitals based on a scheme for partitioning the correlation energy. Many other popular DC type methods, e.g. molecular fractionation with conjugated caps,<sup>20,21</sup> systematic molecular fragmentation,<sup>22-24</sup> generalized energy-based fragmentation,<sup>25-28</sup> are proposed in the literature. In all of these methods, a spatially extended supersystem is divided into smaller, overlapping subsystems (fragments). The total structure, energetics and properties of the supersystem are evaluated from the calculations on the fragments. Though all the above mentioned methods are applicable for geometry optimization, Hua *et al*<sup>28</sup> have attempted to extend generalized energy-based fragmentation method for predicting infrared (IR) spectra of large molecules. This work<sup>28</sup> reports a few prototype computations of Hessian and IR frequencies along with Raman intensities for (H<sub>2</sub>O)<sub>28</sub> and (glycine)<sub>12</sub>. All the calculations reported by Hua *et al*<sup>28</sup> are carried out at HF level of theory and these results show that frequencies and intensities are in fair agreement with their *actual* counterparts, obtained by a full calculation on the supermolecule.

The MTA method developed in our laboratory over the last twelve years is based on a principle in set theory termed as the set *inclusion and exclusion principle*. According to this principle, the *cardinality* of a union of  $N$  overlapping finite sets is defined in terms of the sum of the *cardinalities* of these sets corrected for those of the corresponding overlapping sets. The term *cardinality* means the number of elements in the set: the number of atoms, molecules

and/or bonds in the present context of cluster calculations. This simple principle can be borrowed to partition large, spatially extended chemical systems for exploring their structures, energetics and properties. The idea of Cardinality Guided Molecular Tailoring Approach (CG-MTA, for brevity, referred to as MTA henceforth in the present work) is to fragment the large parent molecule into small overlapping fragments and perform the corresponding *ab initio* calculations only on the fragments. The elegance of the method lies in the use of set inclusion and exclusion principle for obtaining the required property of the parent molecule from the fragments. The MTA method for single point energy calculation and geometry optimization has been discussed in detail in references 7 and 29. The code is developed in line with various other sequential as well as parallel *ab initio* codes developed earlier in our laboratory.<sup>30-32</sup>

The earlier works<sup>6,7,33-35</sup> based on MTA have reported applications of DC-type algorithm for calculating one-electron properties such as Molecular Electrostatic Potential (MESP) and geometry optimization of large molecules. MTA has been recently developed further<sup>36</sup> and benchmarked for the Hessian and frequency calculation of large molecules at HF, DFT and MP2 levels of theory. Li *et al* have done similar attempt<sup>28</sup> for predicting IR normal modes along with respective intensities.

In that work they have demonstrated their method with systems made up of regular units such as polyglycine and water cluster. As their method viz. generalized energy-based fragmentation requires manual fragmentation, application of the method to spatially extended or intricate molecules is difficult. However, MTA uses automatic fragmentation, which easily enables breaking any type of large molecule. Moreover, Li *et al* have demonstrated all the calculations at HF level of theory. On the other hand, MTA-based frequencies can be calculated at HF, DFT and MP2 levels of theory.<sup>36</sup> Currently the development of MTA-based normal mode intensities is underway.

In the present work, we intend to bring out the power of MTA towards a theoretical investigation of large atomic/molecular clusters with a few newly studied examples and several case studies chosen from our earlier investigations. This includes building up several possible structures for clusters, their geometry optimization and study of relative stability of these structures based on their energetics. The MTA has been successfully applied to structurally

diverse metallic and molecular nanoclusters. These studies include  $(\text{CO}_2)_n$ ,<sup>37</sup> orthoboric acid molecular nanoclusters<sup>38</sup> and lithium metal clusters.<sup>39</sup> Along with that, we report some cluster properties evaluated via MTA. The details of the energy and property evaluation are briefly summarized in the section 2.

## 2. Methodology

The present paper reports several case studies detailing the exploration of the structures, energetics and properties of atomic and molecular clusters employing MTA. It is reported in the literature<sup>40</sup> that the number of local minima in potential energy surface increases exponentially with the size of a cluster. Hence, it is very cumbersome and expensive to generate good-quality initial structures for exploring the energetic landscapes. It is found profitable to employ electrostatic guidelines for generating initial structures for molecular and metallic clusters. In our previous works, an electrostatic-based algorithm, viz. Electrostatic Potential for Intermolecular Complexation (EPIC)<sup>41,42</sup> has been developed and exploited for building molecular aggregates. In such an approach, the local minima in the multi-dimensional potential energy surface are located through a general electrostatic-based genetic algorithm. For the best cluster a local optimization within the EPIC framework is carried out employing a direct function minimization Subroutine, STEPIT.<sup>43</sup> The electrostatic interaction energy within EPIC model is defined through a general equation,

$$E_{\text{EPIC}} = 1/2 \left\{ \sum_i V_{A,i} q_{B,i} + \sum_j V_{B,j} q_{A,j} \right\}. \quad (1)$$

In (1),  $V_A$  and  $V_B$  represent respectively the MESP's of the interacting species  $A$ ,  $B$  and  $q_A$ ,  $q_B$  are the corresponding MESP-fitted charges. The MESP at a point  $\mathbf{r}$  is defined as

$$V(\mathbf{r}) = \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} - \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r'. \quad (2)$$

Similarly, for metal clusters, an MESP minimum of the parent cluster acts as a harbinger for the attack of the added metal atom, which is placed in the vicinity of this MESP minimum.<sup>39</sup> These clusters are further subjected to MTA-based geometry optimization as discussed below.

In MTA, the whole cluster is divided into a set of  $k$  main fragments  $\{f_i, f_j \dots f_k\}$ , with secondary over-

lapping fragments  $\{f_i \cap f_j\}$ , and tertiary ones, viz.  $\{f_i \cap f_j \cap f_k\}$ , etc. For this purpose, the best fragmentation scheme for each cluster is carried out based on R-goodness parameter. The details regarding R-goodness parameter are given elsewhere.<sup>7,29</sup> The energy for the whole cluster is calculated from the fragment energies  $E^{f_i}, E^{f_j}, \dots, E^{f_i \cap f_j}, \dots, E^{f_i \cap f_j \cap f_k}$  using the general energy expression<sup>7</sup>

$$E = \sum E^{f_i} - \sum E^{f_i \cap f_j} + \dots + (-1)^{k-1} \sum E^{f_i \cap f_j \cap \dots \cap f_k}. \quad (3)$$

The energy gradients with respect to atomic coordinates ( $\mu$ ) can also be obtained in a similar fashion via the general expression

$$\frac{\partial E}{\partial X_\mu} = \sum \frac{\partial E^{f_i}}{\partial X_\mu^{f_i}} - \sum \frac{\partial E^{f_i \cap f_j}}{\partial X_\mu^{f_i \cap f_j}} + \dots + (-1)^{k-1} \sum \frac{\partial E^{f_i \cap f_j \cap \dots \cap f_k}}{\partial X_\mu^{f_i \cap f_j \cap \dots \cap f_k}}. \quad (4)$$

These energy gradients are fed into the geometry optimization subroutine in the GAMESS<sup>16</sup> package and optimization carried out till the gradient norms become acceptably small. The generated clusters are visualized employing the indigenously developed visualization packages,<sup>44</sup> UNIVIS-2000 and MeTASudio.

The evaluation of the one-electron property (such as MESP) of a spatially extended cluster, over a large, three-dimensional grid of points could be expensive. Hence, following the MTA strategy, the cluster is fragmented in to smaller ones with sufficient overlaps at junction and the property is evaluated from the density matrix generated by picking the appropriate fragment density matrix elements. The MESP is then evaluated using INDPROP<sup>45</sup> code developed in our laboratory. The one-electron property  $P$ , of larger clusters can alternatively be generated via MTA as per the equation

$$P = \sum P^{f_i} - \sum P^{f_i \cap f_j} + \dots + (-1)^{k-1} \sum P^{f_i \cap f_j \cap \dots \cap f_k}. \quad (5)$$

This method of evaluation of MESP of larger clusters is seen to show a good agreement with the corresponding actual one.

More recently, we have calculated MTA-based adiabatic ionization potentials and polarizability of

$\text{Li}_n$  clusters. These values for  $\text{Li}_n$  clusters up to  $n = 22$  showed fairly good agreement with their experimental counterparts. The method has been applied to  $\text{Li}_{40}$  and  $\text{Li}_{58}$  clusters (see ref. 39 for details).

### 3. Structures and energetics of clusters

In the present work, various capabilities of MTA are demonstrated. This section gives an overview of the structure and energetics of a few atomic/molecular clusters, followed by a discussion of various MTA-based properties for these clusters (cf. section 4). In this section, an electrostatics-guided method for building large metal aggregates is demonstrated with the example of lithium clusters. Along with that an EPIC-based algorithm for building  $\text{CO}_2$  prototype clusters is also discussed. These initial structures are subsequently subjected to *ab initio* MTA geometry optimization. This MTA-based geometry optimization method is also illustrated for prohibitively large nanotubes of orthoboric acid, water and benzene molecular clusters, and is discussed sequentially.

#### 3.1 $\text{Li}_n$ metallic clusters

An electrostatics-guided method for building a metal cluster is demonstrated via a prototype  $\text{Li}_n$  cluster. Lithium, being an electropositive atom, the most negative region of MESP of the initial small cluster acts as a harbinger most probable site for the ensuing Li atom. Thus, larger metal clusters are generated using MESP of the smaller motifs as illustrated in figure 1. This general electrostatics-based method to build larger  $\text{Li}_n$  clusters going from  $n = 4$  to 58 is discussed in ref. 39. These smaller clusters are optimized using conventional method, while larger ones are optimized through MTA. The calculations presented below are performed at B3LYP/6-31 + G(*d*) level of theory. This method could predict all the geometries which are reported in the literature. Further in this work, we could build larger clusters, employing electrostatics guidance and MTA optimization. It is observed from our calculations that the larger clusters prefer motifs with more number of hexagonal bipyramids than pentagonal bipyramids. This geometrical transition is justified using molecular electron density (MED) topography.<sup>39</sup> This application of MTA to a metal cluster clearly brings out the possibility of applying such a method in conjunction with a localized basis set rather than popularly used plane wave basis for such investigations.

The MTA-based properties of these lithium metal clusters are discussed in section 4. A similar study on Na clusters is currently under way in our laboratory.

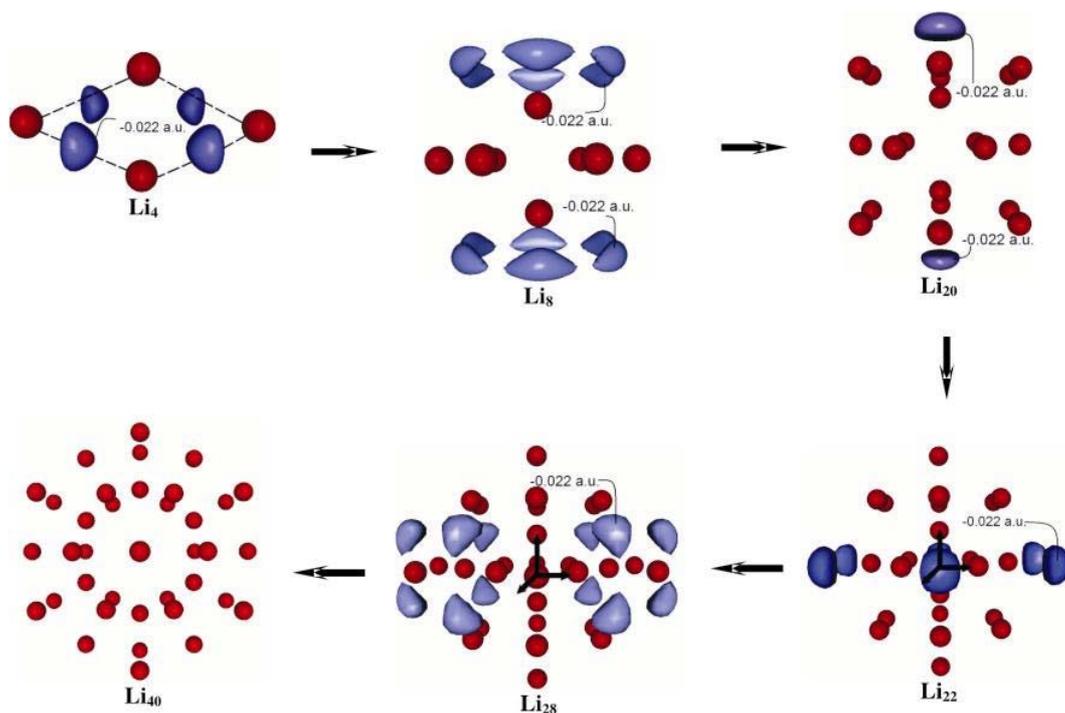
#### 3.2 $(\text{CO}_2)_n$ molecular clusters

The initial geometries of  $\text{CO}_2$  clusters are generated through an algorithm based on molecular electrostatic potential and MTA for building energetically favourable molecular clusters. The details of the algorithm can be found elsewhere.<sup>37</sup> This algorithm is availed for building prototype  $(\text{CO}_2)_n$  clusters with  $n = 13, 20$  and 25. Such clusters are subjected to MTA optimization at MPWB1K/6-31++G(2*d*) level of theory. These MTA optimized clusters with  $n = 13$  and 20 are shown in figure 2. Further, an energy decomposition analysis also performed and which clearly shows that the two- and three-bodies contributes predominantly (>94%) to the total interaction energy. Along with that, we have looked at certain clusters properties, viz. the Hessian, infrared frequencies, MESP and are reported subsequently in section 4.

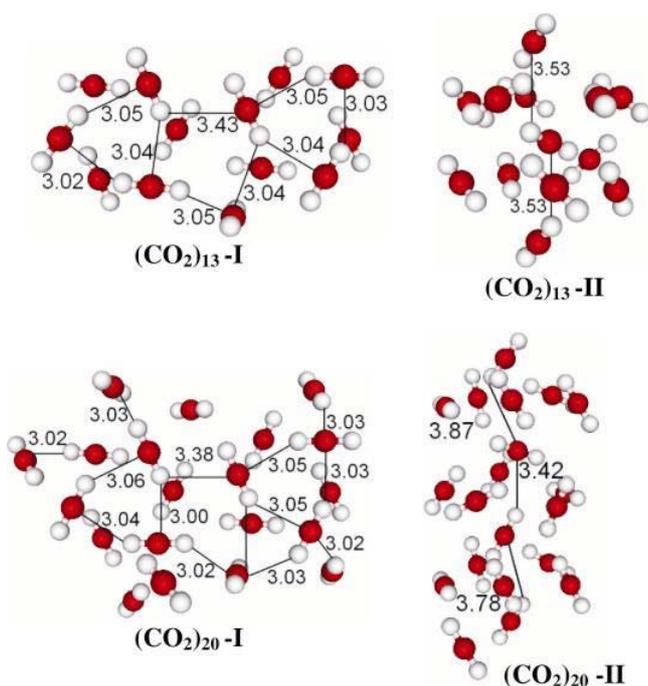
Yet another study on larger  $(\text{CO}_2)_{30}$  clusters at HF/6-31 + G(*d*) level has also been recently attempted. These initial structures are generated through molecular dynamics simulation furnished by Takeguchi,<sup>46</sup> which are subjected to MTA-based optimization. The optimized structures are depicted in figure 3. The distance and angular analysis of the C=O...C interactions is carried out on all the most stable  $(\text{CO}_2)_{30}$  clusters. All these clusters exhibit zigzag patterns similar to that in  $\text{CO}_2$  crystal. The most stable clusters exhibit three interaction spheres with T-type of major interactions. The large clusters are extensively stabilized through C=O...C type of interactions. A critical appraisal of the geometries is done through distance and angular distribution analysis.

#### 3.3 Orthoboric acid nanotubes

Orthoboric acid, viz.  $\text{H}_3\text{BO}_3$ , is a potential building block for nanotubes as each of this unit can form six hydrogen-bonds with other three units. It is evident from the results<sup>38</sup> that nanotubes of boric acid, structurally similar with carbon nanotubes are energetically stable and are stabilized by extensive network of H-bonds. The largest of the calculation reported in reference 38 is that of a nanotube of 64 units,



**Figure 1.** A path of MESP-guided growth of lithium cluster starting from  $\text{Li}_4$  to an  $\text{Li}_{40}$  cluster geometry. The MESP isosurface is shown at  $-0.022$  a.u. See text for details.



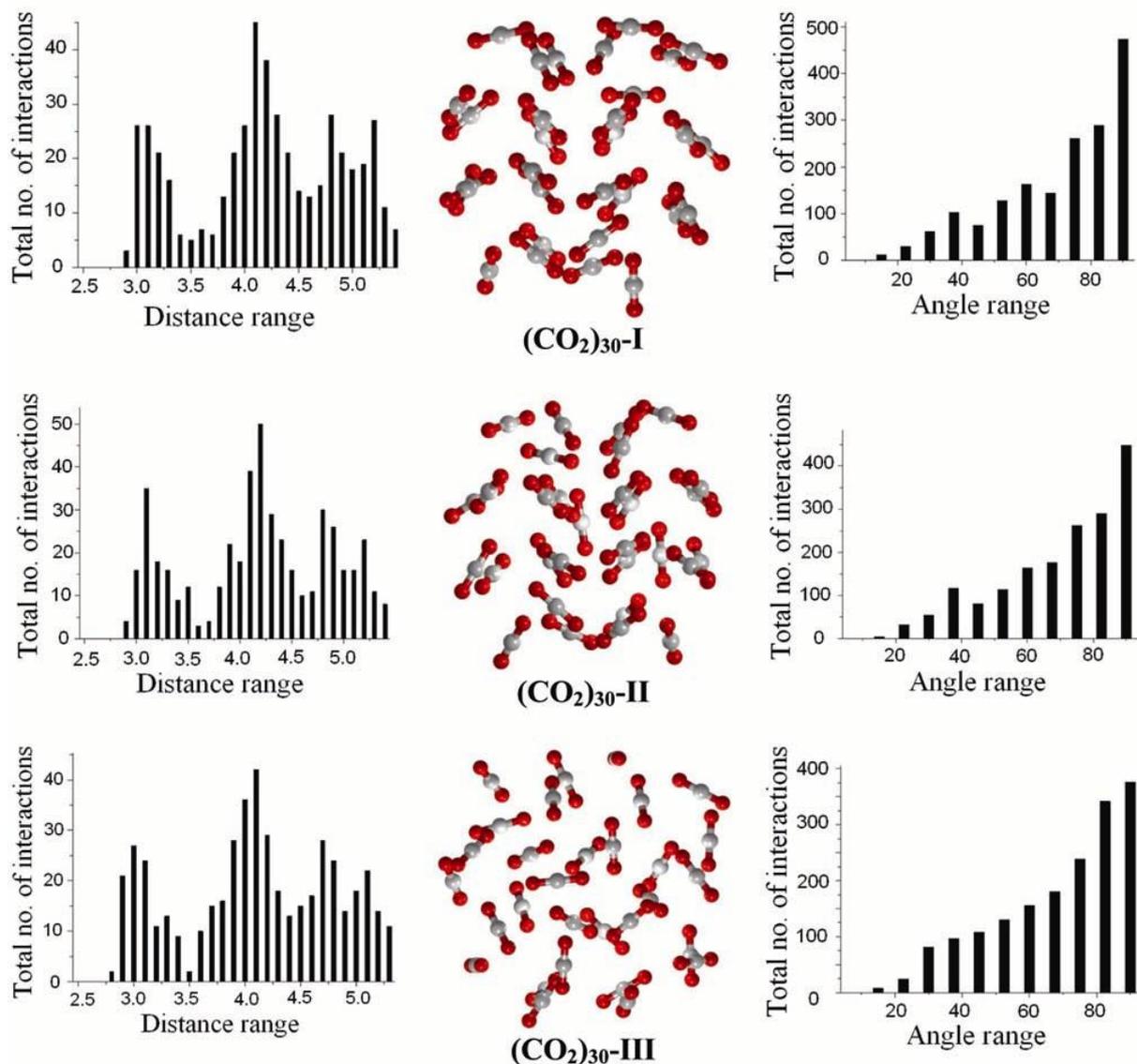
**Figure 2.** MTA-optimized two most stable  $(\text{CO}_2)_n$  clusters each with  $n = 13$  and  $20$  at MPWB1K/6-31++G(2d) level of theory. The distances shown are in angstroms. See text for details.

comprising of 448 atoms (figure 4). This test case of  $(\text{H}_3\text{BO}_3)_{64}$  is presented to highlight the ease with

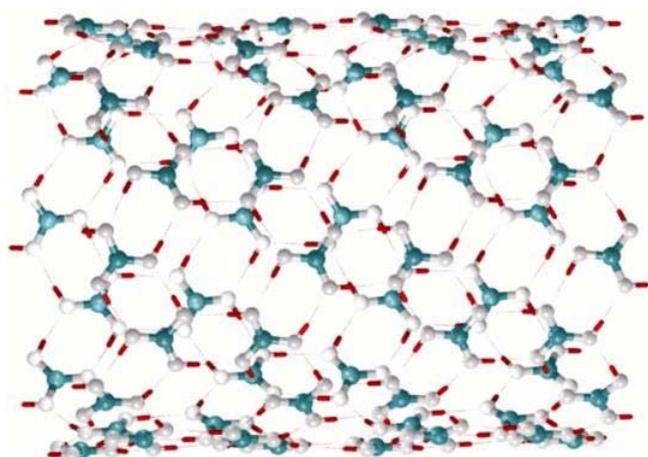
which large molecular clusters can be optimized at *ab initio* level using MTA. This structure is optimized using MTA-enabled GAMESS at HF/6-31++G( $d, p$ ) level of theory. The calculation involves 6016 contractions and hence cannot be performed on PC-based hardware. However, with the aid of MTA, it was performed on a cluster of 8 Pentium IV with 2.8 GHz processor with 2 GB RAM each. The typical time for one SCF cycle was seen to be  $\sim 5$  h on the above mentioned hardware.

#### 3.4 Benzene clusters at MP2 level of theory

The next case study includes benzene tetramers,  $(\text{Bz})_4$  and pentamers,  $(\text{Bz})_5$  which are optimized by MTA-enabled GAMESS at MP2/6-31++G( $d, p$ ) level of theory, in order to show the feasibility of such a large calculation in considerably less amount of time. As the scaling of MP2 method is  $O(N^5)$ , each of these calculations involving 600 basis functions for  $(\text{Bz})_4$  and 750 for  $(\text{Bz})_5$  requires huge amount of memory and computational time and are reported in table 1. The typical memory requirement for  $(\text{Bz})_5$  at MP2/6-31++G( $d, p$ ) is 32 GB, while the MTA calculation requires at most 5 GB. Thus in such cases, MTA brings down the computational



**Figure 3.** The distance analysis, geometry and angular analysis on the MTA optimized three most stable  $(\text{CO}_2)_{30}$  clusters at HF/6-31+G(d) level of theory. See text for details.



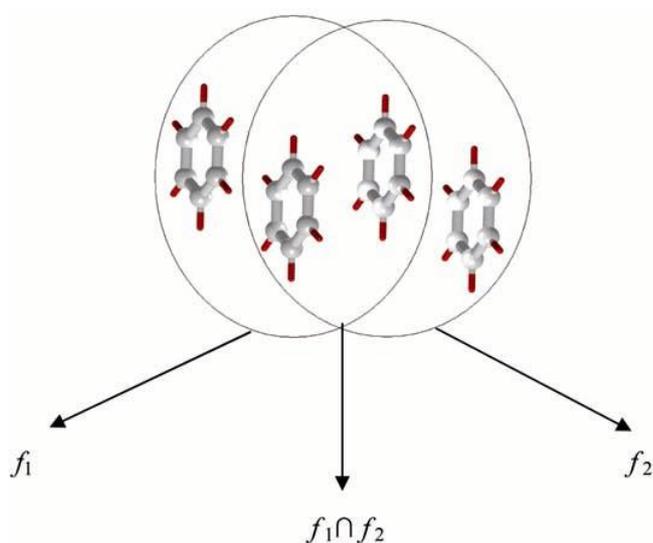
**Figure 4.** MTA optimized  $(\text{H}_3\text{BO}_3)_{64}$  cluster at HF/6-31++G(d, p) level of theory. See text for details.

cost and time. Geometries of these clusters optimized at MP2/6-31G(d) level of theory are taken as the initial geometries for this study employing higher basis 6-31++G(d, p). For each of the tetramers and pentamers, 2 to 3 trimer fragments can be cut so as to cover all the important interactions. A prototype fragmentation scheme is illustrated in figure 5. MTA-based energy and gradients for the final optimized geometries (figure 6) are reported in table 1, along with the time taken per cycle of optimization on a single Intel Core 2 Quad at 2.4 GHz processor with 8 GB RAM and 250 GB of disk. The current study shows that the motifs with T-type interactions are energetically more favourable over shifted-parallel and parallel stacked ones.

**Table 1.** Benzene tetramers and pentamers: MTA estimates of energy,  $E$  (a.u.), interaction energy,<sup>a</sup>  $\Delta E$  (kcal/mol), maximum and root mean square gradients, ( $G_{\text{MAX}}$ ,  $G_{\text{RMS}}$ ) (a.u.), CPU time,  $t$  (min/optimization-cycle) and number of main fragments ( $N_{\text{F}}$ ) used for MTA-based geometry optimization. See text for details.

Geometry	$E$	$\Delta E$	$G_{\text{MAX}}$ , $G_{\text{RMS}}$	$t$	$N_{\text{F}}$
(Bz) <sub>4</sub> -I	-926.11129	-20.4	0.0009, 0.0002	520	4
(Bz) <sub>4</sub> -II	-926.11616	-23.5	0.0002, 0.0000	540	4
(Bz) <sub>4</sub> -III	-926.12155	-26.9	0.0009, 0.0002	240	2
(Bz) <sub>5</sub> -I	-1157.63661	-23.9	0.0002, 0.0000	212	2
(Bz) <sub>5</sub> -II	-1157.64128	-26.9	0.0019, 0.0004	370	3
(Bz) <sub>5</sub> -III	-1157.64072	-29.5	0.0017, 0.0004	366	3

<sup>a</sup>Energy of benzene monomer optimized at the same level is -231.51969 a.u., is used for calculating the interaction energy



**Figure 5.** Schematic fragmentation of a benzene tetramer [(Bz)<sub>4</sub>-I].  $f_1$ ,  $f_2$  are the main fragments, while  $f_1 \cap f_2$  is the overlapping fragment of  $f_1$  and  $f_2$ . See text for details.

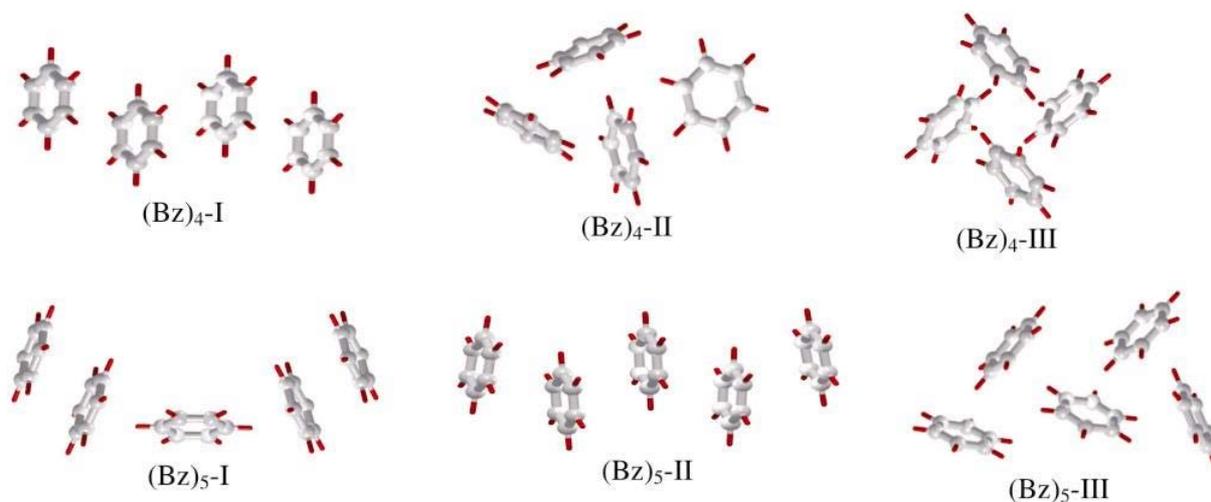
## 4. Properties of clusters

### 4.1 Hessian and IR frequencies

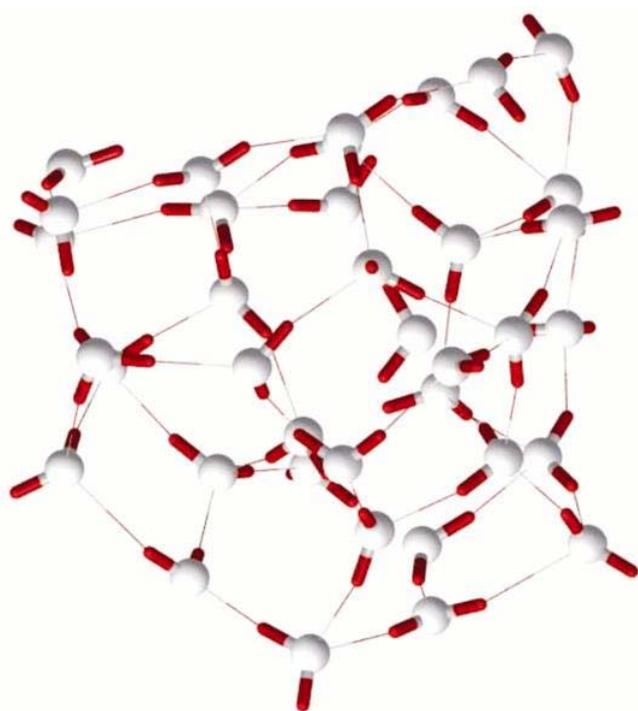
MTA has been extended<sup>36</sup> to evaluate Hessian for a macromolecule from its fragment-Hessians, and thus, it also enables to calculate the frequencies of the macromolecule. Here we present a small test case of (H<sub>2</sub>O)<sub>37</sub> (refer figure 7 for structure) for MTA-based Hessian and frequency calculation at HF/6-31G. This calculation involves 481 contractions, and hence, it is possible to perform actual calculation in order to demonstrate the accuracy of the method. However, large test cases, where it is not possible to perform frequency calculation with any

hardware that is available today, are reported in ref. 36. Both the actual and MTA-based calculations on (H<sub>2</sub>O)<sub>37</sub> are performed on a cluster of 4 Intel Core 2 Quad processors with 2.4 GHz, each with 4 GB RAM and 250 GB hard disk. The MTA-based calculation took merely 4 h and 30 min while, on identical hardware, the actual calculation took ~20 h. This shows that our method has notable time advantage over the conventional one. In terms of accuracy of the MTA-Hessian as compared to actual one, the maximum error is  $2.9 \times 10^{-3}$  a.u. and root mean square deviation (RMSD) is 0.07. The only significant frequencies for water cluster are O-H stretches. A wide range of these from 3171 to 4106 cm<sup>-1</sup> is obtained, which is due to extensive hydrogen bonding between the water molecules. Percent maximum error and RMSD in frequencies are 0.48 and 0.028 (in cm<sup>-1</sup>) respectively. Thus, it is clear that MTA enables computation of Hessian and vibrational frequencies without significant loss of accuracy saving considerable amount of computational time.

The above test case establishes the reliability of the method. Therefore, a test case where actual calculation is not at all possible is presented next.<sup>36</sup> A nanotube of forty (H<sub>3</sub>BO<sub>3</sub>) units is subjected to such a MTA-based Hessian and frequency calculation at HF/6-31++G( $d$ ,  $p$ ) level of theory. The calculation involves 280 atoms and 3760 basis functions. For (H<sub>3</sub>BO<sub>3</sub>)<sub>40</sub>, the frequencies within range of 3970–4190 cm<sup>-1</sup> correspond to O-H stretches, while those in range 1120–1600 cm<sup>-1</sup> are due to B-O stretches. The calculation is run on a cluster of 13 dual core machines with 2 GB RAM each. It is worth pointing out that Hessian is expensive to calculate and such a calculation for 3760 basis functions can not be performed on commonly available hardware with con-



**Figure 6.** MTA-optimized structures of three most stable benzene tetramers and pentamers each at MP2/6-31++G(*d*, *p*) level of theory. The respective energies, interaction energies and gradients are reported in table 1. See text for details.



**Figure 7.** MTA optimized  $(\text{H}_2\text{O})_{37}$  cluster at HF/6-31G(*d*) level of theory. See text for details.

ventional methods. It is noteworthy that the corresponding MTA-based calculation merely took 12 h. Other details about the fragmentation scheme used for this nanotube are elaborated in ref. 36.

Vibrational frequencies of  $(\text{CO}_2)_n$  clusters with  $n = 13$  and 20 are evaluated through MTA. This cal-

culcation is performed at MPWB1K/6-31++G(2*d*) level of theory, involving 975 and 1500 basis functions for  $(\text{CO}_2)_{13}$  and  $(\text{CO}_2)_{20}$  respectively. For a comparison, MTA calculations are performed on  $(\text{CO}_2)_{13}$  along with the corresponding actual ones. The MTA Hessian matrix elements are found to be in agreement with the corresponding actual ones up to two decimal digits. The most intense asymmetric stretching mode of  $(\text{CO}_2)_{13}$  is at  $2501\text{ cm}^{-1}$ , while the respective MTA value is at  $2502\text{ cm}^{-1}$ . Further, all the MTA-based frequencies show a fair agreement within  $1\text{ cm}^{-1}$  with the actual one. For  $(\text{CO}_2)_{20}$ , actual frequency calculation is cumbersome, hence only MTA-based frequencies are evaluated. The highest asymmetric stretching mode for this cluster is found<sup>37</sup> at  $2509\text{ cm}^{-1}$  and is blue-shifted with respect to  $\text{CO}_2$  monomer at  $2499\text{ cm}^{-1}$ .

#### 4.2 Polarizability, adiabatic ionization potential and MESP

An optical property of interest is the polarizability and this is a measure of the distortion of the delocalized valence electron in the presence of an external electric field. This study for  $\text{Li}_{20}$  is reported to exemplify the accuracy of MTA method. MTA-based polarizability values of  $\text{Li}_{20}$  are 1238, 1244 and 1249 a.u. with three different fragmentation schemes with increasing R-goodness parameter.<sup>7</sup> This MTA-based polarizability is found to be in good agreement with the actual calculated value at 1263 a.u.

An important electronic property of interest to chemist is adiabatic ionization potential (AIP). The adiabatic ionization potential is defined as the respective energy difference of the neutral and cationic clusters. The larger cationic cluster is also optimized imposing +1 charge to the respective fragments. The difference of the neutral and charged systems optimized with same fragmentation scheme is reported as AIP of larger clusters. The details of the fragmentation scheme and values will be reported elsewhere in the literature.<sup>39</sup> In this work,<sup>39</sup> the AIP for  $\text{Li}_n$  clusters with  $n = 2$  to 18 are evaluated by conventional calculation and that for larger clusters with  $n = 20$  to 58 are calculated with MTA. The MTA-based (experimental) AIP for  $\text{Li}_{20}$ ,  $\text{Li}_{22}$ ,  $\text{Li}_{24}$  are 3.37 (3.31), 3.57 (3.39) and 3.49 (3.24) eV respectively. The errors in MTA-based AIPs are typically 2–7% in comparison with their experimental counterparts. With similar assumptions, we have calculated<sup>39</sup> AIP of  $\text{Li}_{40}$  and  $\text{Li}_{58}$  to be 3.43 and 3.24 eV respectively, showing a decreasing trend in AIP values with increasing  $n$  for larger clusters.

Molecular electrostatic potential is also evaluated for these large clusters through MTA. The larger clusters are fragmented into smaller ones and the MESP is evaluated over the grid points around these fragments. The respective fragments MESP over the grid point is patched up based on the general MTA expression. The MTA-based MESP's of these clusters show a remarkably good agreement with the corresponding actual ones. A critical comparison of this property evaluated via both the ways is illustrated in ref. 37. It is noted that the MESP minimum and the function distribution around the molecule is qualitatively and quantitatively similar to the actual one.

## 5. Concluding remarks

The present work has given a brief account of various capabilities of MTA for exploring the structure, energetics and properties of large clusters. Initially, a concise outline of MTA including the fragmentation and optimization procedure is presented. From our earlier studies, it is evident that electrostatic guidelines are useful for generating good initial geometries for atomic/molecular clusters. Electrostatic coupled with MTA-based optimization can be employed to locate local minima on the PES of molecular or metallic clusters. This is illustrated with the examples of  $(\text{CO}_2)_n$  and  $\text{Li}_n$  clusters in the

present work. The above mentioned test cases as well as orthoboric acid nanotubes and benzene clusters elucidate the power of MTA to perform such huge calculations merely on a PC-based hardware, which are conventionally impossible. The level of theory is no bar for this approach. It can be applied at various levels of theory with an equal ease and this is demonstrated in the current work through a wide spectrum of test cases at HF, DFT (for MPWB1K and B3LYP functionals) and MP2 levels of theory.

Further, MTA can be exploited for studying variety of properties of large clusters. These properties include ionization potential, polarizability, Hessian and IR-normal mode frequencies. The ionization potential of Li-clusters is in good agreement (error <7%) with their actual counterparts, proving the reliability of this method. The numerically significant Hessian elements are seen to match up to three digits after the decimal point. Further, the IR frequencies are found to agree within  $1 \text{ cm}^{-1}$  and these values on comparison with the actual show a fair agreement. Altogether MTA is a good model to understand the structure, energetics and properties of large clusters in different domains.

It may be pointed out that the current version of MTA is restricted to closed shell systems and there is a scope for implementing it for open shell, excited state and transition state structure and property evaluations. This will enable further applications of the method to chemically and physically meaningful systems. The future outlook of the method is promising. We anticipate employing MTA for investigations of very large clusters which can eventually be correlated with crystal structures as well as surface characteristics of materials in the coming five years. We hope that in the longer term it may also be possible to perform on-the-fly molecular dynamics calculations with such an approach.

## Acknowledgements

Authors are grateful to the Council for Scientific and Industrial Research (CSIR), New Delhi and Center for Development of Advanced Computing (C-DAC), Pune. SRG thanks the Department of Science and Technology (DST), New Delhi for the award of J C Bose Fellowship. We are thankful to Dr G Narahari Sastry for furnishing initial geometries of benzene clusters and to Dr H Takeguchi for the MD-based geometries of  $(\text{CO}_2)_{30}$  clusters.

## References

- Baletto F and Ferrando R 2005 *Rev. Mod. Phys.* **77** 1 and references therein
- Crocker J C 2008 *Nature* **451** 528
- Spangler D and Christoffersen R E 1972 *Proc. Nat. Acad. Sci. USA* **69** 3301
- Yang W 1991 *Phys. Rev. Lett.* **66** 982
- Yang W and Lee T S 1995 *J. Chem. Phys.* **103** 5674
- Gadre S R, Shirsat R N and Limaye A C 1994 *J. Phys. Chem.* **98** 9165
- Ganesh V, Dongare R K, Balanarayan P and Gadre S R 2006 *J. Chem. Phys.* **125** 104109
- Nakano T, Kaminuma T, Saro T, Akiyama Y, Uebayasi M and Kitaura K 2000 *Chem. Phys. Lett.* **318** 614
- Kitaura K, Sugiki S, Nakano T, Akiyama Y and Uebayasi M 2001 *Chem. Phys. Lett.* **336** 163
- Fedorov D G and Kitaura K 2004 *J. Chem. Phys.* **121** 2483
- Fedorov D G, Olsen R M, Kitaura K, Gordon M S and Koseki S 2004 *J. Comput. Chem.* **25** 872
- Fukuzawa K, Kitaura K, Uebayasi M, Nakata K, Kaminuma T and Nakano T 2005 *J. Comput. Chem.* **26** 1
- Fedorov D G, Ishida T, Uebayasi M and Kitaura K 2007 *J. Chem. Phys.* **A111** 2722
- Fedorov D G and Kitaura K 2007 *J. Phys. Chem. A* **111** 6904
- Fedorov D G, Jensen J H, Deka R C and Kitaura K 2008 *J. Phys. Chem.* **A112** 11808
- The GAMESS package, Schmidt M W, Baldrige K K, Boatz J A, Elbert S T, Gordon M S, Jensen J H, Koseki S, Matsunaga N, Nguyen K A, Su S J, Windus T L, Dupuis M and Montgomery J A 1993 *J. Comput. Chem.* **14** 1347
- Kobayashi M, Akama T and Nakai H 2006 *J. Chem. Phys.* **125** 204106
- Kobayashi M and Nakai H 2006 *Chem. Phys. Lett.* **420** 250
- Kobayashi M, Imamura Y and Nakai H 2007 *J. Chem. Phys.* **127** 74103
- Zhang D W and Zhang J Z H 2003 *J. Chem. Phys.* **119** 3599
- Chen X, Zhang Y and Zhang J Z H 2005 *J. Chem. Phys.* **122** 184105
- Deev V and Collins M A 2005 *J. Chem. Phys.* **122** 154102
- Collins M A and Deev V 2006 *J. Chem. Phys.* **125** 104104
- Netzloff H M and Collins M A 2007 *J. Chem. Phys.* **127** 134113
- Li W and Li S 2004 *J. Chem. Phys.* **121** 6649
- Li S, Li W and Fang T 2005 *J. Am. Chem. Soc.* **127** 7215
- Li W, Li S and Jiang Y 2007 *J. Phys. Chem.* **A111** 2193
- Hua W, Fang T, Li W, Yu J G and Li S 2008 *J. Phys. Chem.* **A112** 10864
- Gadre S R, Rahalkar A P and Ganesh V 2006 *IANCAS Bull.* **4** 267
- Gadre S R, Kulkarni S A, Limaye A C and Shirsat R N 1991 *Z. Phys. D-Atoms, Molec. and Clusters* **18** 357
- Gadre S R, Limaye A C and Shirsat R N 1993 *J. Comp. Chem.* **14** 445
- Gadre S R and Limaye A C 1993 *J. Chem. Phys.* **100** 1303
- Babu K and Gadre S R 2003 *J. Comput. Chem.* **24** 484
- Babu K, Ganesh V, Gadre S R and Ghermani N-E 2004 *Theoret. Chem. Accts.* **111** 255
- Gadre S R, Babu K and Ganesh V 2005 *Recent trends in practice and theory of information technology: Proceedings of NRB Seminar* (ed.) S N Maheshwari (New Delhi: Viva Books)
- Rahalkar A P, Ganesh V and Gadre S R 2008 *J. Chem. Phys.* **129** 234101
- Jovan J K V and Gadre S R 2009 *Int. J. Quant. Chem.* **109** 2238
- Elango M, Subramanian V, Rahalkar A P, Gadre S R and Sathyamurthy N 2008 *J. Phys. Chem.* **A112** 7699
- Jovan J K V and Gadre S R 2008 *J. Chem. Phys.* **129** 164314
- Malliet J, Boutin A and Fuchs A H 1998 *J. Chem. Phys.* **109** 329
- Pundlik S S and Gadre S R 1997 *J. Phys. Chem.* **B101** 3298
- For the original work and applications of EPIC model, see Pundlik S S and Gadre S R 1997 *J. Phys. Chem.* **B101** 9657; Pingale S S, Gadre S R and Bartolotti L J 1998 *J. Phys. Chem.* **A102**, 9987 and Shivasen D, Babu K and Gadre S R 2000 *J. Phys. Chem.* **A104** 10887
- The subroutine STEPIT, written by Chandler J P 1965 Physics Department, Indiana University, distributed by QCPE, University of Indiana, Bloomington, IN
- (a) UNIVIS 2000: A comprehensive visualization package, Limaye A C and Gadre S R 2001 *Curr. Sci. (India)* **80** 1296; (b) Ganesh V *J. Comput. Chem. (ASAP)* MeTASstudio: Available at: <http://code.google.com/p/metastudio>
- (a) INDPROP, the molecular property calculation package developed by Theoretical Chemistry Group, Department of Chemistry, University of Pune, Pune (India). See: Shirsat R N, Bapat S V and Gadre S R 1992 *Chem. Phys. Lett.* **200** 373; (b) Balanarayan P and Gadre S R 2003 *J. Chem. Phys.* **119** 5037
- Takeuchi H 2008 *J. Phys. Chem.* **A112** 7492